DOCUMENT RESUME

ED 211 585 TM 820 007

AUTHOR Alkin, Marvin C.; And Others
TITLE Evaluation Use Project. Final Report of the User Interview Survey.
INSTITUTION California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY National Inst. of Education (ED), Washington, D.C.
PUB DATE Nov 81
GRANT NIE-G-80-0122
NOTE 197p.

EDRS PRICE MF01/PC08 Plus Postage.
DESCRIPTORS Administrators; *Decision Making; Elementary Education; *Evaluation; Interviews; *Research Utilization; School Personnel; *Surveys; *Use Studies
IDENTIFIERS *Evaluation Utilization

ABSTRACT

The User Interview Survey was conducted by the Evaluation Use Project (EUP). The goals of the survey, which were achieved, were gathering, categorizing, and analyzing information about evaluation use among elementary school decision makers. Most survey respondents agreed instructional and curricular issues were most important. It was found that most school decision makers did not frequently rely upon evaluations when they made decisions. Needs assessment evaluations were most helpful to school staff in identifying areas requiring attention. Evaluative data were used more frequently in curricular decisions and in those involving the bilingual program. The data did not usually enter into administrative, staff development, or personnel decisions. The level of evaluation utilization increased in those decisions in which administrators participated. Refinements in the observations and data collection procedures directed toward underlying causes were suggested. The implications for future research on evaluation utilization and evaluation practice in the schools were discussed. The analyses conducted as part of this research have the potential to increase evaluation utilization at the local level. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Center for the Study of Evaluation

UCLA Graduate School of Education
Los Angeles, California 90024

ED211585

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

CSE

CSECSE

CSECSECSE

CSECSE

CSE

TM 820 007

DELIVERABLE - November, 1981

EVALUATION USE PROJECT

Final Report of The User Interview Survey

Marvin C. Alkin

Project Director

Grant Number

NIE-G-80-0122

P-4

CENTER FOR THE STUDY OF EVALUATION

Graduate School of Education

University of California - Los Angeles

The project presented or reported herein was performed pursuant to a grant from the National Institue of Education, U.S. Department of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by the National Institute of Education should be inferred.

Final Report of the User Interview Survery

Brian Stecher, Marvin Alkin

&

Gretchen Flesher

November, 1981

© Regents of the University of California

Acknowledgements

We would like to acknowledge the assistance of Ross Connor, whose insightful comments greatly improved the final document. We would also like to thank Richard Daillak and Frederica Geiger for their assistance with various sections of the analysis. Finally, we must recognize the invaluable participation of our Research Team: Eloise Appel, Donna Davis, Jonathon Horowitz, Phyllis Jacobson, Paula Stern, Steven Stumpf and Peter White, who participated in the data collection and conceptual analysis phases of this study.

CONTENTS

Chapter		*page*
1.	INTRODUCTION	1
	Precedents to CSE Evaluation Use Research	2
	Prior Work of the Evaluation Use Project	5
	Evaluation Case Studies.	5
	Evaluator Field Study	7
	The User Interview Survey	8
2.	METHODOLOGY	11
	Interview Strategy	11
	Interview Format	13
	The Topic-Centered Interview.	14
	Selection and Training of Interviewers	16
	The Research Team	16
	Interviewer Validator Training	17
	Selection of the Respondent Sample	20
	Defining the Population	20
	Contacting the School District	21
	The Field Interviews	23
	Scheduling Interviews	23
	Conducting Interviews	24
	Data Analysis	25
	Developing the Initial Written Data Base	26
	Validating the Initial Written Data Base	27
	Conceptual Data Synthesis	31
	Instrumental Data Synthesis	36
	Quantitative Analysis	42
3.	DESCRIPTIVE STATISTICS: DECISION MAKING PHASE	47
	Introduction	47
	Breakdown of Significant Occurrences	47
	Discussion	54
	Breakdown of Types of Information	59
	Discussion	63
	Personnel Configurations in School-Level Decision Making	65
	Discussion	69
4.	INTERRELATIONSHIPS: DECISON MAKING PHASE	71
	Introduction	71
	Decision Type vs Information Type	72

Incidences of Increased Evaluation Use. 74
Incidences of Decreased Evaluation Use. 78
Other Observations. 78
Discussion 81
Personnel Configuration vs. Information Type 82
Discussion 88
The Relationships Among Information Types 90
Discussion 90
Information Type vs. Other Variables 91
Discussion 100

5. DECISION PATTERNS ACROSS ALL FOUR PHASES 103

Introduction 103
Analysis of the Decision Phases 103
Comparison between the Phases 103
Recognition 106
Ratification 108
Dissemination 109
Discussion 109
Prototype Sequences (by Decision Type) 110
Discussion 118

6. CONCLUSIONS 121

Summary 121
Refinements 127
Implications 129
Research Implications 129
Practical Implications 132

BIBLIOGRAPHY 137

LIST OF TABLES

Table		*page*
1.	Type of Significant Occurrence	54
2.	Information Sources	63
3.	Personnel Configurations	68
4.	Frequency of Information Use in Each Phase	104
5.	Personnel Configuration in Each Phase	105

LIST OF FIGURES

Figure *page*

1. Information Use: All Decisions, N=73 73
2. Information use: GEN CURR(N=6) vs. TOTAL(N=73) 76
3. Information use: BILINGL(N=7) vs. TOTAL(N=73) 77
4. Information Use: INS MATL(N=13) vs. TOTAL(N=73) 80
5. Information Use: TEACHERS(N=53) vs. NO TEACH(N=20) . . . 84
6. Information Use: NO ADMIN(28) vs LO ADMIN(31) vs HI ADMIN(14) 86
7. Information Use: NO T+ADM(N=43) vs. T+AD GPS(N=30) . . . 87
8. Information Use: ONE OPTN(31) vs FEW OPTS(10) vs UNKNOWN(32) 93
9. Information Use: SHORT(31)vs MEDIUM(32) vs LONG(10) . . . 95
10. Information Use: PRINCIPL(22) vs GROUP(13) vs NO STRAT(25) 96
11. Information Use: INT PRMP(N=37) vs EXT PRMP(N=33) 98
12. Information Use: ADMINIST(20) vs TEACH GP(9) vs MIXED GP(6) 99

Chapter 1

INTRODUCTION

The User Interview Survey was undertaken to understand how evaluation information is used, how much it is used, by whom it is used, and under what social/institutional/political conditions it is used. The User Survey, the third study on evaluation utilization conducted by the Evaluation Use Project (EUP) at the Center for the Study of Evaluation at UCLA, examined patterns of information use among elementary school decision makers. Our goal was a better understanding of mix of evaluation and other information inputs into program decisions and of the relationship between information and decision making.

The User Survey was a logical successor to the two earlier studies conducted by the Evaluation Use Project -- the Evaluation Case Studies (Alkin, Daillak & White, 1979) and the Evaluator Field Study (Daillak, 1980). In addition, it owes some debt in its formulation to the accumulated knowledge concerning evaluation utilization derived from a variety of research studies over the past decade. A full understanding of the genesis of the User Interview Survey and the importance of the results requires some knowledge of the historical background of evaluation utilization research and the previous efforts of the Evaluation Use Project.

PRECEDENTS TO CSE EVALUATION USE RESEARCH

The field of evaluation grew to prominence in the late 1960's with the increased federal commitment to social welfare programs. The Elementary and Secondary Education Act (ESEA) of 1965 and other legislation required that program evaluations be conducted annually. In fact, often specific sums of money were earmarked for this purpose.

Such rapid growth in the amount of evaluation being done allowed little time for a systematic assessment of its impact on program decision making. When the first assessments emerged, they were quite pessimistic. Writers complained about the quality of evaluation and its consequent lack of impact. Guba summarized what he viewed to be "the most obvious clinical signs of evaluation's failure".

> Any professional area that is so much avoided; that produces so many anxieties; that immobilizes the very people who want to avail themselves of it; that is incapable of operational definition, even by its most trained advocates, who in fact render bad advice to practitioners who consult them; which is not effective in answering reasonable and important questions, and which has made little apparent effort to isolate and ameliorate its most serious problems -- must indeed give us pause. (p. 31).

There is little wonder, given such an assessment, that evaluation was seen to have little impact on decision making. However, it should be noted that such widely accepted judgments -- however stridently offered -- generally were not the result of empirical research; they were based primarily on what might be labeled 'informally-shared personal experience.'

Several writers speculated on factors that explained this limited use of evaluation information. Arson & Sherwood (1967) commented upon the importance of diplomacy and rapport. Reviewing the course of one

evaluation they noted a number of areas of conflict between evaluators, program designers and practitioners and concluded that "skill in the craft (of evaluation) requires more than technical knowledge. In fact, the ability to be diplomatic is perhaps as important as any. (p. 96)

Mann (1972) underscored the importance of proper methodology. He reviewed 181 evaluation reports and found that they did not even meet the minimum requirements for technical quality. He concluded that mistakes of the kind found throughout the sample are "extremely damaging to the cause of evaluative research. With two or three exceptions, the errors are of a major character. In other areas of research in the behavioral sciences, any of them would probably render a study unfit for publication." (p. 275)

Rodman & Kolodny (1964) focused more on organizational factors, basing their exposition on personal experience as well as a review of other writings. They discussed the importance of work and time organization, patterns of communication and other related factors in the structure of the agency being studied and how these affect the use of evaluation research.

Weiss (1966) called for systematic study of the impact of evaluation, but none was undertaken until the mid 1970's. Two prominent studies of the period, Alkin et al. (1974) and Patton et al. (1975), used systematic survey research techniques to carry out their investigations.

Alkin and his associates studied the impact of evaluation on decision making in a sample of 42 ESEA Title VII programs at both federal and local levels. At the federal level they found that program evaluation had little perceptible influence on decision making -- just as the earlier

literature had suggested. They found no relationship between evaluation reports and funding levels or federal monitor's ratings of project quality. At the local level, however, quite different findings emerged. Project directors reported that evaluations *had* affected their decisions to modify their programs during the year and had assisted them in other important areas as well.

Patton and his colleagues looked at 20 health care programs and their evaluations. They also found that evaluation *did* have an impact, but not in "organization-shaking" ways. Instead, evaluation tended to provide "additional information" helpful to program decision makers and considered by them, though not always the most important consideration.

An important consequence of Patton's research was a heightened awareness of the importance of subjective, interpersonal factors in evaluation utilization, in addition to structural and systematic variables. Patton specifically asked about 11 factors commonly identified in the literature as affecting utilization. Of these, only one, the political factor, was deemed important by his informants. However, an entirely new and different factor emerged as *the* most important influence on utilization: "the personal factor". This factor involved the attitudes, interest, abilities and actions of key decision makers. As Patton explained,

> Utilization is not simply determined by some configuration of abstract factors; it is determined in large part by real, live, caring human beings. (p.37)

Taken together, these two studies suggest that the earlier writings had overlooked some important aspects of utilization. Program managers

and decision makers *close* to the evaluation -- not distant recipients of an anonymous and impersonal evaluation report -- may be the most likely users of the evaluation information. Moreover, the uses to which the information is put may be incremental and low-key rather than dramatic go/no-go decisions about program continuance. One major consequence of this research was that the earlier, narrow conception of evaluation utilization came under attack. As Patton pointed out:

> The results of our interviews suggest that what is typically characterized as under-utilization or non-utilization of evaluation research can be attributed in substantial degree to a definition of utilization that is too narrow and fails to take into consideration the nature of actual decision-making processes in most programs. (p.10)

Thus, by the late 1970s evaluation utilization was recognized as a dynamic, incremental process in which the discretionary actions of individual evaluators or decision makers influence the ultimate disposition of an evaluation's findings as much as -- and perhaps more than -- the political and organizational features of the system.

PRIOR WORK OF THE EVALUATION USE PROJECT

Evaluation Case Studies.

The research of the mid-1970's pointed out that evaluation utilization was a subtle and complex process. The goal of the EUP over the past three years has been to develop as complete a picture of evaluation utilization as possible. We first tried to depict these subtleties more clearly, using qualitative, naturalistic methods. Five in-depth case studies of Title I or Title IV-C school programs and their evaluations were undertaken. Using open-ended interviews and extensive field observations, Alkin, Daillak & White (1979) constructed a detailed description of program implementation and evaluation at each school.

Based on these case studies, Alkin et al. developed a framework for the study of utilization which identified the major personal and contextual factors to be considered at the local level. (See Appendix A.) Many of the dimensions that emerged were familiar, though portrayed in richer detail than before. The study captured vividly the complexities of local decision making, and illustrated the cumulative, incremental nature of the utilization process. The study also highlights the importance to utilization of the expectations and attitudes of the decision makers, a finding that corresponds to the "personal factor" identified earlier by Patton. However, the most potent element observed by Alkin and his colleagues was not the personal characteristics of the decision maker, but rather the personal characteristics of the *evaluator*. The use of a "consultative" approach by the evaluator appeared to have greater potential for increasing utilization than any other element identified in the study.

The Evaluation Case Studies suggested several approaches that an evaluator might take to increase the impact of evaluation at the local level. Some elements identified were beyond the evaluators' control; others -- especially those related to evaluation approach -- could be purposely manipulated. In the case studies, local program managers had responded positively to evaluators who took an adoptive, "helper" or "user-focused" approach. However, the case studies had not focused on the wider organizational structures within school districts that could constrain possible evaluator roles. The whole issue of the circumstances of the evaluator had not been addressed.

Evaluator Field Study

The Evaluator Field Study (Daillak, 1980) addressed the evaluator's professional position and organizational situation. Daillak spent a year accompanying three evaluators in Metro district as they conducted their various activities. As participant observer, he witnessed their interactions with colleagues, the district administration and local school personnel. He saw the impact that bureaucratic structures had on their scope of action, as well as the impact of personal relationships, resource constraints, attitudes and expectations.

Daillak concluded that there were strong organizational impediments to useful evaluation in the schools. Local school administrators were generally disinterested in, or even hostile to, evaluation. Informal discussions of test results and other evaluative information were possible between evaluator and school administrator, but anything beyond that was shied away from by both parties. In this manner the bureaucracy effectively limited the formal role of the evaluator. The evaluation consultants, as the evaluators were called in Metro district, were channeled into reporting and technical assistance functions, and there was no real opportunity to assume a consultative role in their official capacity.

But the evaluation consultants supplemented their reported work through informal, unreported contacts. In this manner some evaluators could promote more "planful" instruction despite the strictures of their official bureaucratized role. Thus, while the school organizational structure effectively circumscribed the classical evaluator role, the creative evaluation consultant went outside official channels and adopted

an approach that is in line with the recommendations one might draw from our earlier research.

THE USER INTERVIEW SURVEY

From our earlier research it was clear that evaluation information was just one of many possible inputs into decision making, and that the evaluator was one person among many who interacted with the school administrative staff. The narrow focus on evaluation and evaluators produced an unbalanced picture of evaluation's impact on school decisions by highlighting the occasions when evaluation did come into play and spotlighting the personnel who were directly concerned. It would have been premature to formulate recommendations without knowing more about the competing inputs and actors in the decision making process. Those concerns were addressed in the User Interview Survey.

To understand the role evaluation played in program decisions, the EUP needed to look at a broad cross-section of significant program decisions and consider all the elements involved in the process, including -- if relevant -- evaluation and evaluators. The reality was that program-related decisions were being made all the time at each school. Input to these decisions came from a variety of sources, only some of which could be considered to be evaluation. The key personnel in these decisions included the site-level administrators, classroom teachers and parents, as well as evaluators. In fact, as the Evaluator Field Study suggested, the evaluation personnel had only intermittant impact.

The goals of the User Interview Study, then, were to obtain a better understanding of the significant areas of school decision making, to ascertain the relative importance of evaluation in these school decisions, and to determine what role might realistically be projected for evaluation. The methodology employed to accomplish this task will be described in Chapter 2; the results of the study will be presented in Chapters 3, 4 and 5.

Chapter 2

METHODOLOGY

The User Interview Survey sought to place evaluation's contribution to the school's working environment in an appropriate context. The interviews would explore the issues central to the daily concerns of school administrators along with the various sources of information that were relevant to their decision making. In this chapter we describe the elaboration of the interview strategy and format, the selection of the respondent sample, the training of the interviewers, the collection of the interview data, and the analytic procedures that were employed with these data.

INTERVIEW STRATEGY

To obtain the local informants' point of view, the interviewers solicited the respondents' portrayal of some significant recent occurrences in the program and of the salient forces or considerations that affected these occurrences. After hearing the informants' account of these matters, the interviewer probed for what evaluation did or did not contribute to these events. This procedure provided a local perspective on which activities were considered significant and worthy of special attention and on how local decision maker's responses were formulated. We learned both *who* was involved in the occurrence and *what* the basis was for their actions.

After exploring evaluation's connections, if any, with these salient interviewee concerns, the interviewer asked about the primary emphasis of the program's recent evaluations and about the impact these evaluations had had upon the school program. Finally, if time permitted, the discussion was shifted from concrete events and circumstances to more general issues, which permitted the interviewee to expand a bit on his or her attitudes toward evaluation in general -- its usefulness, and its problems.

There was an underlying rationale which guided the adoption of this strategy. If the interview had opened with direct questions about evaluation, it might have had the effect of "leading the witness" to overstate the importance of the issues treated in the evaluation and of evaluation's significance to program operation. Instead, the interviewee should identify specific significant program occurrences first and discuss recent evaluations later. By grounding the interview in specifics, the survey hoped to escape the generalities and platitudes that might be expected in an abstract discussion of evaluation's virtues, faults, and impacts.

The interview probed "significant occurrences" rather than, for example, "significant decisions" or "significant concerns" following the argument so skillfully made by Weiss (1980). Weiss argues that in bureaucratic organizations policy actions often are not "decided" but rather "accrete" in a gradual flow of "small uncoordinated steps taken in many offices -- by staffs who have little awareness of the policy direction that is being promoted or the alternatives that are being foreclosed" (p. 382). A "significant occurrence in the life of the

program" was more tangible and more likely to be something informants at each school could recognize, discuss, and analyze than the narrow "decision" or negative "concern". It connoted a change or departure from the ordinary stream of activity in the school -- an opportunity for influence, something that evaluation might (or might not) have affected.

INTERVIEW FORMAT

Two basic concerns guided the choice of interview method. First, the interviews needed to elicit, with sensitivity and a minimum of distortion, the respondent's accounts of the "who's" and "what's" of significant occurrences in their programs. Second, the EUP nevertheless, had an agenda of specific interests to explore with the interviewees. We considered a number of possible formats that imposed varying degrees of structure on the interview, and selected the procedure that best satisfied these two concerns. A short digression will help explain this choice.

Harold Levine, at UCLA, offers what he terms the Questionnaire-Jawboning Continuum as a useful construct for thinking about the use of structure in data collection. At the questionnaire extreme, the data exchange is totally structured. Respondents answer only the questions asked, with only the answers provided. The data collector has no opportunity to tailor the interaction to the individual respondent. While such a data collection strategy offers tremendous comparability across subjects, its sensitivity is limited to the choices built into the instrument. "Jawboning" defines the other extreme of the research continuum: A nearly unstructured conversation between two persons,

without a specific agenda or external structure. Jawboning can be rich in detail and sensitive to subtle ideas and nuances of meaning, but "jawboning" data lacks comparability from subject to subject.

Between the two extremes, there is a variety of data collection options. For example, questions can be carefully structured and sequenced, but the interviewee can be allowed open-ended responses. Alternately, an interviewer might be allowed to conduct a seemingly free-flowing conversation with the subject, after which the interviewer might complete a very structured, forced-choice questionnaire reporting on the interaction.

The Topic-Centered Interview.

Initially, we considered using a structured interview format with subjects being allowed open-ended responses, but rejected this choice as too rigid to capture the diverse range of stories we expected to hear from our respondents. In its place, we selected what we termed a "topic-centered" interview format. Such a format places a modest amount of structure on the interviewer -- by outlining in a "topic guide" the topics to be covered in the interview -- but leaves specific questions and probes to the discretion of the interviewer. The respondent is almost unfettered, except as the interviewer may take steps to refocus the respondent's remarks or move the discussion along to other topics. Thus, the topic-centered interview offers great flexibility within a guiding framework.

Patton (1980) discusses much the same method in his description of the use of an "interview guide":

> An interview guide is a list of questions or issues that are to be explored in the course of an interview. An interview guide is prepared in order to make sure that basically the same information is obtained from a number of people by covering the same material. The interview guide provides topics or subject areas within which the interviewer is free to explore, probe, and ask questions that will elucidate and illuminate that particular subject. Thus, the interviewer remains free to build a conversation within a particular subject area, to word questions spontaneously, and to establish a conversational style -- but with the focus on a particular subject that has been predetermined. (p. 200)

The study's topic guide (or interview guide, as Patton would have it) is displayed in Appendix B. It implements the overall interview strategy in a manner that is sensitive to both of our initial concerns. The specific research topics are identified and form the framework for the conversation. Within this framework the interviewers are free to explore the respondents' ideas fully and with a minimum of distortion.

The brevity, indeed the almost skeletal quality of the guide, underlines the key ramification of using such a format: interviewer training *must* be comprehensive and thorough. The training with its supporting materials (See Interview Survey of Users: Interim Report, 1980, and Appendices C & D.) inculcates in the interviewers the rationale and purpose of the interviews; explains in exhaustive detail the kinds of information which should be sought out under each topic; and prepares the interviewers for the verbal interaction they must establish successfully to secure meaningful, high-quality data. The guide, then, becomes simply a set of cues to the interviewers, helping them recall the elements of their training.

SELECTION AND TRAINING OF INTERVIEWERS

The Research Team

The interviewers were drawn from a group of advanced graduate students in Research Methods and Evaluation at UCLA enrolled in a graduate seminar on Evaluation Utilization in the spring of 1980. All students participated in a five-week training sequence. Interviewers were selected by the end of the third week, with the remaining trainees selected as validators. (The role of validators will be discussed in a subsequent section.) In the last two training sessions we were able to divide into subgroups and have the interviewers practice interviewing the validators, who role-played school personnel.

The selection of interviewers was based on a number of factors. First, it was important that the interviewer have some direct school experience. Actual work in a school setting for an extended period of time gave our interviewers a background for understanding nuances and subtleties of school-related decision making and provided a knowledge framework within which to pose questions.

Second, we wanted to select, based on the principal investigator's observation, those trainees displaying the highest general maturity and interpersonal skills and the greatest interviewing skill. Most of the group performed at a high level on all dimensions; both interviewers and validators were actually very well qualified. On these bases, five interviewers and five validators were selected.

Interviewer Validator Training

Staff training involved four phases: (1) understanding of the evaluation utilization research; (2) training in general interview skills; (3) familiarity with the District Administration, organizational structure, and terminology; and (4) familiarity with the specifics of this particular project and its procedures.

To ensure that research team members all had a reasonably competent understanding of relevant literature in evaluation utilization, all read and discussed *Using Evaluations: Does Evaluation Make a Difference?* (Alkin et al., 1979). All had read Michael Q. Patton, *Utilization-focused Evaluation* (1978) as part of an earlier training session. In addition, trainees read other articles on utilization, including major pieces by Carol Weiss, Nathan Caplan, Larry Braskamp et al., and Jane David.

The interview training sessions were conducted by Harold Levine, Department of Education, UCLA, who is an anthropologist and expert on interviewing, and by Marvin Alkin. Their presentations involved lectures, videotapes of model interviews, discussions, practice interviews, and reactions. During these activities, both Levine and Alkin observed the trainees and noted those who were mastering the interview strategies most effectively.

To familiarize the trainees with the context in which this research would be conducted, Richard Daillak gave a presentation about the organization of the Metro Evaluation and Testing Office, the activities commonly engaged in by the evaluators, and the kinds of assessment commonly found in the schools. A glossary of common school terms,

particularly as related to special programs and their evaluation, was presented to each trainee.

Finally, a number of materials specifically related to the detailed procedures of the study were developed. The topic guide (See Appendix B.) was a one-page summary of the main topic headings and areas of interest, and was designed for use by interviewers in the actual interview situations. The interview topic description contained explanatory materials on the meaning and scope of the different interview dimensions. (See Appendix C.) A mock interview narrative consisted of a complete facsimile transcript of the interviewer portion of an interview. The data reporting and summary forms will be described in greater detail in a subsequent section on data aggregation.

These training materials were developed by the senior members of the research team. Other members of the research team and outside experts reviewed and modified the various training materials during their development.[1] In addition, pilot interviews at an eligible school tested the research framework and the interview topic guide procedures. These interviews proved to be quite useful in refining and properly targeting the training materials. Based on the field tests and other reviews, the senior researcher revised the order in which the interview topics were presented and modified the suggested phrasing of questions.

---

[1] Two colleagues merit special thanks. Carol Weiss, who reviewed some of the materials during a visit to UCLA, and Michael Patton, who stood ready by phone and mail. Both provided characteristically generous and perceptive advice. We are glad to acknowledge their superb assistance. Naturally, though, they bear no liability for the final product: that is ours alone.

When sufficient familiarity with the topic guide and with interview techniques had been achieved, trainees also reviewed and discussed transcripts and tapes from the pilot interviews. It became apparent, for example, that keeping the interview "on-target" constituted one of the most difficult tasks; the school personnel's narrative tended to ramble and wander. Often their free-flowing monologues did uncover valuable insights, but from time to time it was necessary to refocus the discussion. As a part of their training, interviewers compiled a valuable repertoire of conversation-directing probes and phrases.

Each interviewer went through one additional hour-long simulation of the complete interview sequence from entering to leaving the decision maker's office. A member of the study team played the role of school decision maker, mimicking the cooperative, but often disorganized, responses that had been encountered in the pilot interviews. During the interview, the surrogate decision maker took notes about the interviewer's questions, successful and unsuccessful strategies, and content material which the interviewer had failed to obtain with his or her particular questioning. After the conclusion of the interview, the two discussed the experience in detail and the "decision maker" suggested areas for improvement.

A further phase of the training occurred after the first school interview had been conducted. One interview tape was selected; the research team listened to the tape together and each person summarized the conversations on the data summary form to be discussed in a subsequent section. Comments about the summary forms were elicited, and during the discussion that followed, some minor modifications in the

forms were made. More importantly, however, these exchanges allowed the research team to standardize each person's interpretation of how to summarize conversations, what certain topic descriptions should contain, and what certain questions meant.

SELECTION OF THE RESPONDENT SAMPLE

Defining the Population

The respondent sample of school site level decision makers was determined largely by the research circumstances. Since the User Interview Study was intended to complement Daillak's (1980) field study of evaluators in "Metro District", Metro school staff needed to be included. Metro, a large urban district, could be adequately covered only by multiple interviews. Field interviews are labor-intensive and project resources were limited. The almost inescapable consequence, therefore, was to limit the study to the Metro district.

The Evaluation Use Project's historical concern with the evaluation of specially-funded programs naturally directed our attention to "school site level decision makers" connected with such programs. More important, Metro did little program evaluation except of specially-funded programs. (Actually, Metro typifies many school districts in this concentration of program evaluation activity.) We decided to limit this study by focusing on schools receiving Title I funding, first, because one could be sure such schools had experienced evaluations (since Title I requires them) and second, because the program offers a large pool of schools from which to sample.

We decided to interview three individuals at each school site, in part to obtain overlapping responses from multiple informants to "triangulate" our data (in the jargon of qualitative research) but also in part because Metro's Title I schools seem to have multiple important "decision makers" (Daillak, 1980). The school principal was interviewed in every case. In addition, two other persons holding influential, knowledgeable positions relating to the school's programs were selected. Such positions have a number of different names; based on our previous contact with the school system, we developed a working list of all acceptable job titles. As one of the two additional interviewees, we selected a person who had specific coordinating responsibility for the special program. (This person was usually called "Title I Coordinator" or "Special Program Coordinator" or "Assistant Principal".) The final respondent at each school was a staff person who was involved in some manner in the administration of the special program. In a large school, there might be an individual whose job was entirely administrative. In a small school, it was often necessary to include people with the title of "Resource Teacher", "Curriculum Supervisor" or "Bilingual Coordinator".

Contacting the School District

The Superintendent of Metro agreed that the project was worthwhile, committed the District's participation, and directed the Evaluation and Testing Office (E & T) to assist in sample selection. Nonetheless, participation on the part of individual schools was voluntary, and we anticipated that some schools would be reluctant to give the time neces-

sary to participate in the research. For this reason we overselected schools. We asked for a preliminary random sample of 28 schools from the much larger population of all Title I elementary schools though we planned to conduct interviews in only 20 schools. The district compiled the desired sample, which included schools from all geographic areas of the district as well as schools of diverse size and ethnic composition. Each school principal received a letter from the Superintendent which briefly described the study, endorsed its purposes, and vouched for the researcher's credentials -- but also established that school participation was completely voluntary. (A copy of this letter and other study materials will be found in a prior report; Alkin, Stecher & Daillak, 1980.)

In follow-up telephone conversations, all but two principals expressed a willingness to participate, and we halted sample selection once 20 principals had committed themselves and their schools to the study. Then we augmented this sample with two additional schools, selected from those serviced by the compensatory education evaluator studied by Daillak (1980) in the companion study to this research. Thus, 22 schools ultimately participated in the study.[2]

As stated, almost all the principals agreed to participate: only two declined. One school principal asked to be excused because "participation was voluntary". She added that she was without an assistant principal, had additional duties, and needed to give any extra time she might have to the children and teachers. The second principal also

[2] Later, one school dropped itself from the sample after its principal became ill and was unable to participate in the interviews. That school was replaced by another school selected randomly from the preliminary sample of 28 candidate school sites.

mentioned that participation was voluntary. He declined to participate because this was his first year as principal of the school and all his time and energy was needed to help solve existing problems within the school. Both principals were very apologetic and wished us well on the project.[3]

THE FIELD INTERVIEWS

Scheduling Interviews

Two or three weeks before the field interviewing was to begin, research team members called each principal and explained the proposed interview procedures. They arranged for three one-hour interviews with 10 or 15-minute breaks between interviews. Principals were asked to identify two other members of their staff who were school level decision makers as we had defined them earlier.

Though this identification procedure was not random, we doubt that it introduced any bias into our results. This is because it was only in the case of the respondent that the principal exercised any significant amount of free choice. Most schools did have more than one additional individual with administrative responsibility who fit our criteria for the third person. But even here the principal's selection criteria (whatever they were) had little bearing on that person's ability to recall events, and hence had little impact on the generalizability of our results. There was little or no flexibility in the selection of the first two respondents: the principal was always interviewed, and we always

---

[3] The remaining non-selected schools in the preliminary sample were contacted by phone and letter, thanking them for their cooperation but informing them that the randomly-selected final sample was filled.

asked to speak with the Title I Coordinator, if such a person existed. If no Title I coordinator existed, we asked for the highest ranking administrator with responsibility for Title I programs.

We tried to obtain the name and the official title of each of the interviewees in our first phone call. If the names and titles of the other two school level decision makers had not been obtained in the initial call, they were obtained during a second telephone confirmation, a day or two before the interview. In one or two instances, the scheduled interviewee was unavailable when the interviewer arrived at the school, and the principal had arranged for a substitute who satisfied our respondent selection criteria. Almost without exception, the school personnel we dealt with were cooperative and willing to go out of their way to meet our requirements.

Conducting Interviews

The interviews were conducted without major problems. The first interview always was conducted with the school principal, and, before it began, the rest of the day's schedule was reconfirmed. In addition, the interviewers generally secured, in advance, an appropriate location for each interview. We thought it important that the interviews not take place in a public place; not only could distractions interfere with the conversation, but respondents also might find it difficult to answer candidly while their peers were within earshot.

Each of the interviews was tape recorded on identical machines. (Since tape counters are not standardized from one brand to another, identical tape recorders facilitated subsequent data analysis and review.)

Tape recordings were overt. At the beginning of each conversation, the interviewer indicated that he was planning to make a taped record of the interview to ensure accuracy in the study and to facilitate future analysis. There were no objections to this, although a few of the respondents asked that the machines be turned off momentarily while they made certain comments. In each of these isolated instances, the respondent commented about another individual at the school and did not want the person's name recorded.[4] Aside from these instances in which personalities were involved, there were no irregularities or surprises in the inteview process.

DATA ANALYSIS

The aggregation of field data is one of the most difficult tasks for those who conduct naturalistic research. Hours of interviews and pages of notes must be summarized systematically into a usable form. A balance must be struck between maintaining the richness of detail afforded by the naturalistic data and reducing data sets to a manageable and comparable form. A number of procedures have been tried by different researchers to accomplish this task. Alkin, Daillak and White (1979) presented a multi-stage data aggregation strategy in *Using Evaluations*. The strategy used in the current study is guided by that approach, while at the same time it varies from some specific procedures because of the nature of the data.

---

[4] We indicated to each respondent that all data would be recorded anonymously at the beginning of the interview, but such assurances are not always remembered...or believed. In fact, one of the subtle disadvantages of tape recording is that voices are identifiable, and the actual tapes themselves are never truly anonymous.

The data analysis proceeded through several phases: developing and validating an initial written data base, undertaking various first stage data synthesis activities, refining the data base, and carrying out the final stage analysis and synthesis. Each of the procedures will be discussed in the sections which follow.

Developing the Initial Written Data Base

In our view, one of the most critical points in the analyses of qualitative data is the development of the initial written data base. Knowledge and insights gained from previous research enabled us to focus the interviews on five specific topic areas. This simplified data aggregation by providing a logical framework within which interview and validation summaries would be fitted. Summary forms were developed corresponding to the initial topic guide. (See Appendix D for the interviewer summary form.)

As soon as possible after conducting the three interviews, the interviewer set about the task of completing a summary form for each interview. Respondents were coded by school (e.g., 17) and by position (e.g., SP2--the second staff person interviewed). The first step was to summarize accurately the actual information conveyed by the respondent. Interviewers referred to their notes of the conversation as they recorded comments within each of the topic areas of the summary form. The second step was to listen to the tape to select direct quotations which captured the significant information and perspectives embodied in

---

[5] Though the process was not overly complicated, it was nonetheless quite time consuming. The summary form for each one-hour interview took two and one-half to three hours to complete.

the comments of each respondent.[5] In addition, the interviewers elaborated on their initial written summaries if the tape recording suggested important details they had omitted. Thus, the final summary form contained five or six pages of narrative comment on the respondent's point of view (organized by topic area) and up to an additional five or six pages in relevant direct quotations. The summary form, along with a second summary form to be discussed shortly, became the initial data base for subsequent analysis.

Validating the Initial Written Data Base

A number of strategies was employed to ensure the accuracy of the initial written interview summaries. In particular, four project procedures helped to assure data validity: use of tape recorders, use of independent validators, internal verification, and external verification.

Tape Recorders. Arguments have been raised against the use of tape recorders, (e.g., they are intrusive, artificial, a mechanical crutch, etc.) However, there are also strong arguments in their favor:

1. they free the interviewer to concentrate more on developing his or her next questions instead of recording the respondent's previous answer and allow the interviewer to focus his or her attention on the respondent rather than a piece of paper;
2. they allow the interviewer to replay the interview and listen for things that might not have been readily apparent during the interview; and
3. they serve as a permanent record of the raw data of the study.

This latter argument, in our view, constitutes the most important reason for using tape recorders. The permanent raw data base allowed us to secure a second, independent written summary of each interview and thus provided a means to validate the interviewer's impressions. Later, after we narrowed our analytic focus and developed final coding procedures, we reassessed the raw data tapes a third time. Such a thorough, multistage analysis would have been impossible without this permanent record.

Independent Validators. After an interview had been conducted at a school and the summary forms completed by the interviewer, the cassette tapes were turned over to a validator. Working from the tapes alone, this person completed a second independent set of summaries. Validators listened to the tapes (and completed their summary forms) in the order in which the interviews took place. Each tape was played completely through before the validators began the process of summarizing the interview according to the topics in the validator's summary topic guide. (Validator's summary forms paralleled those used by the interviewers.) A second listening of the tape generally produced the remainder of the information necessary for the summary forms. Validators also identified and transcribed key quotations from each respondent. Frequently this required listening to the tape a third time; occasionally only portions of the tape needed to be reviewed.

Internal Verification. The two summaries together (interviewer's and validator's) provide the basis for within-project verification of the accuracy of the initial written data base. A step-by-step comparison was made of each pair of summary documents. A high correspondence would

allay fears that the data aggregation process might have introduced individual biases or discontinuities.

We had anticipated the possibility of substantial discrepancies between the two versions of each interview and had developed a procedure for adjudicating these differences. A panel, consisting of the interviewer, the validator, and a third member of the research team, would consider both written versions of the interview and, if necessary, would listen to the interview tape before ascertaining the correct interpretation of the actual events.

In fact, while some differences between the validator and interviewer summary sheets existed, the differences were (almost without exception) in the amount of detail included while reporting the same occurrence or point of view. After the initial comparisons, there were only two or three instances in all 65 hours of tapes in which the interviewer and the validator reported information which was contradictory. Moreover, none of these discrepancies centered on a focal issue in the interview. Relistening to a portion of the tape recordings provided a simple but satisfactory resolution of differences. As a result, we are confident that our data aggregation process accurately portrayed the interview information.

External Verification. It is also possible that what was actually said during the interview did not accurately reflect the respondent's point of view, perhaps because of the interview content. The interviewer, for example, frustrated or distracted the respondent with repeated interruptions to ask for clarification or additional detail. We already knew that our summaries accurately reflected what *had been said*. External

validation would tell us if what had been said accurately portrayed the situations and points of view of the respondents.

A sample of respondents was asked to judge the accuracy of the summaries of their own interviews. The second school visited by each of the five interviewers was selected for field validation.[6] Copies of the interviewer summary forms were mailed to the three respondents at each of these five schools. They were asked to review the summaries and note inaccuracies. We asked them, "Do these summaries accurately reflect the events you described?" This field validation process, therefore, gave us a measure of the sensitivity of our interviews. We learned if the words that were said accurately portrayed the situations and points of view of the respondents.

Follow-up phone calls were made a week after the mailing, reminding respondents to return the summaries with their comments. The close of the school year precluded a second set of reminders. Nevertheless, 10 of the 15 summaries were returned. (It is our belief that a respondent who found errors in the summary was more likely to return it than one who felt everything had been portrayed accurately.) Four of the ten respondents made corrections. A total of 26 comments were made on the other six forms.

An analysis of the respondents' comments revealed very few substantive differences with the summaries. In most cases, elaborations and explanations offered by the respondents represented tangential information that had not come out in the interview process. In sum, detailed analysis of the comments affirmed that our interpretations of events and

---

[6] We felt that the first set of interviews would not be truly representative. (Nor would the last one.)

respondents' points of view were quite valid.

Conceptual Data Synthesis

Though we had specific objectives for the User Interview Study and a number of research questions for which we sought answers, we approached the analysis cautiously. One problem with analyzing qualitative data is that the researcher tends to impose his or her own categories rather than letting the data "speak for themselves". To avoid this pitfall we began the data synthesis in a non-directive manner. In unstructured group discussions we collected impressions and identified areas for further scrutiny.

As the group discussions progressed, certain themes began to emerge repeatedly from the comments of different respondents, and we focused on these inductively derived topics. We elected to investigate a variety of these themes and developed a procedure called the Human Data Bank to facilitate verification of preliminary notions against the full collection of written summaries. We proceeded further with some of the analyses and produced working papers on a small number of different themes. The underlying relationships that emerged in this manner became the basis for our later structured data synthesis, and the variables that we deemed to be important after our conceptual synthesis were included in those instrumental data refinement activities. The group discussions and the Human Data Bank will be described below.

Group Discussions. The synthesis of data from the data base began with a series of open discussions among the members of the research team -- the principal investigator, five interviewers, and five valida-

tors. The group met weekly following the completion of interviewing and validation to share impressions and experiences. Research team members already were beginning to notice patterns among responses in their data subsets which might hold across the complete sample. Their discussions touched upon many topics, including exactly what constituted "significant occurrences" in the minds of our respondents, what types of data seemed to be the most important to them, and what their reactions were to the different kinds of evaluation data that were available.

After three group meetings, each member of the team was asked to prepare a draft report based on the interviews he/she conducted or validated. Team members were asked to make their report reflect only the data from their own data base. The discussion of these draft reports at a subsequent staff meeting was very enlightening. A surprising number of points of view emerged. One person saw the management style and administrative approach of the respondents as the most significant variable. Another focused on the favorable or unfavorable results of the Program Quality Review (PQR) process. A couple of staff members commented on the wide variety of respondent impressions about what the word "evaluation" actually meant. Some very interesting and useful insights emerged from this discussion.

To obtain an external critique of the themes emerging from the data, a conference telephone call was arranged with Michael Patton. Members of the research team discussed their initial thoughts with Patton, a process which resulted in a good many insightful and illuminating comments.

Descriptive Analysis and The Human Data Bank. We continued our conceptual synthesis of the data in two other forms. First we categorized and summarized a number of prominent features of the respondents in our sample and the significant occurrences they described. For example, job titles of respondents were categorized and similarly, an initial coding system for type of significant occurrence was developed, and a breakdown of significant occurrences was produced. These summary descriptions helped us to familiarize ourselves with the massive data base we had gathered. In addition they suggested a number of interesting initial patterns for further analysis.

While categorizing and coding features for an inductive summary review was a relatively easy task, a more complex strategy was necessary for the bulk of the analysis. Initially, we thought that once a preliminary topic of interest had been identified in the group discussions, we would search the complete data base of interview and validator summaries for information pertaining to the topic. Unfortunately, case-by-case review was time consuming, and each subsequent review seemed an inefficient duplication of effort. This is both a blessing and a curse of qualitative research: repeated review and examination uncover subtleties and nuances, but it prohibits simply turning quantitative data over to a computer programmer for a quick statistical printout.

We developed a compromise technique. Five of the researchers, those who had been the most extensively involved and had done the greatest number of interviews or validations, wanted to continue working on the project until some of the analyses were completed.

Having this group of researchers available suggested another possible scheme for analysis: "the human data bank". (While the label makes the principal investigator cringe, he has yet to find an acceptable substitute.)

The five research team members and the principal investigator were each responsible for the distinct group of schools with which they were the most familiar. During the following months of analysis they acted as channels to the data from their schools. They reviewed the interviews and both summary sheets in detail and made brief notes to aid in later recall. Team members became informed stand-ins for the actual raw data.

The analysis team met regularly throughout the summer. Each person selected one of the themes which had emerged from our earlier work to pursue in greater detail. For example, if one member had an idea about how the decision maker's personality affected his or her attitude toward the data, this would be presented to the group. Each member would comment on the idea based on the information given by the respondents in his or her group of schools. After such a discussion it was usually easy to tell if a line of inquiry was worthy of further investigation, needed modification, or should be abandoned.

When an idea appeared worthy of further investigation, the person leading that inquiry drew up a questionnaire or a series of direct probes which could be put to the "human data bank". Members of the group prepared a detailed response, identifying specific relevant examples and relating direct quotations from the respondents. Most importantly, they also identified code numbers which could be used to locate

the information in the summary sheets. The human data bank respondents thus served like a card catalogue or index.

Finally, based on input from the human data bank, the analyst of a particular topic reviewed the data summary sheets themselves. After this perusal of the data, drafts of analytic papers were prepared. They included: complete descriptions of the ideas or relationships that were being investigated, a discussion of the data on which they were based, direct quotes to explicate the presentation, and further elaboration to explain moderating elements in the analysis and interactions. Each draft was reviewed by the full team and, in essence, was compared with the data reality as perceived by the "data bankers". Only drafts which stood up to the scrutiny of the complete group were refined and developed into project working papers. Because of the participation of the other members of the group, not only were the papers that evolved at this stage stronger and more thoroughly thought out, but the process of checking them directly against the raw data also was significantly simplified.

We learned a great deal from these conceptual data synthesis efforts. We identified a number of important variables that seemed to be related to evaluation use. These included the types of information that were available, the personal style of the administrators, the number of options or alternative course of action that were considered, whether or not someone "championed" a particular cause, the personality of the district evaluation consultant, and so forth. More importantly we found an overall structure for analyzing the events that had been described to us. Our respondents' descriptions of significant occurrences were

almost all organized around decision making processes. Whether these were imposed decisions from administrative superiors, individual decisions by principals, or deliberative processes carried out among the whole school staff, decision making seemed to be at the core of the occurrence. Further analysis of decision making procedures seemed to hold the greatest potential for probing evaluation utilization in these significant occurrences. To summarize, then, our conceptual data synthesis culminated in the identification of a number of important variables for further study and the emergence of decision making as the core around which to structure that study.

Instrumental Data Synthesis

As noted above, the more our familiarity with the data increased, the more our attention was drawn towards the decision making process as the key structure underlying evaluation utilization in each significant occurrence. Evaluation utilization seemed to be inextricably linked to decision making, and a fuller understanding of the decision process might shed useful light on utilization. In particular, we hoped to be able to characterize patterns in school level decision making and to investigate the role that different information types -- including evaluation -- played in these actions.

One problem emerged, however; though our interviews contained a lot of information on decision making, summaries lacked sufficient detail for such an analysis. Our initial notion had been that the written summaries would provide a sufficient base for all further study. It was only after we progressed sufficiently in the task of data synthesis that

we became aware of their shortcomings: they were broadly focused to convey a valid representation of the whole interview, but they lacked the precise information we desired on this specific topic.

While the broad notion of a "significant occurrence" encompassed many possible school actions, most discussions focused on a key prominent decision that the school made in relation to the occurrence. Thus decision making had been portrayed in considerable detail in most of the interviews. However, obtaining these detailed accounts necessitated developing new instrumentation and relistening to the raw data tapes. This reanalyis procedure is discussed below.

Developing the Coding Form. Our goal in reanalyzing the tapes was to describe the decision-making process that had been elucidated in the interviews in a manner that allowed us to examine patterns in the data and relationships between variables of interest. One objective was to understand the relative importance of evaluative information vis-a-vis other kinds of inputs into elementary school decision making. Another objective was to see if any relationship existed between the type of decision and the range of information brought to bear upon it. To do this we needed a framework for organizing the relevant data from the interviews. A framework corresponding to decision theory seemed logical.

Decision theory suggests that problem analysis proceeds through a number of phases before its ultimate resolution (e.g., Griffiths, 1958). While the number of stages and the identifying labels vary from author to author, all agree that the first phase entails recognition of a problem or need for action. This is followed by a process of interactions among

the parties directly involved in the decision, until ultimately a single course of action is selected.

Our respondents, too, talked about activities at the school proceeding through a sequence of steps that ultimately resulted in some response or action. However, our earlier data synthesis efforts suggested that decision making in the schools does not end with the identification of the ultimate course of action. In fact, there may be two more identifiable activities before the process achieves complete resolution. Many times a recommendation arrived at through a process such as the one we had described was subsequently "ratified" by the principal or by the staff as a whole. While there was the potential for a veto at this stage, more often the selected course of action was given *pro forma* approval. In addition, there was sometimes a follow-up stage in which information about the decision was disseminated to a wider audience -- either the general school staff, advisory bodies, parent organizations or the broader local community.

Consequently, we hypothesized a four-phase model to structure the analysis: (1) identification of a prompt to action; (2) an interaction process culminating in a specific decision; (3) possible review and "sign-off" by other school personnel; and (4) possible dissemination of the decision to a wider audience. At each point in the process, we identified which actors were involved and what kinds of data -- personal impressions, quantitative measures, expert recommendations, etc. -- contributed to their actions. Evaluative data were of particular interest.

Insights gained from our earlier analyses and our current project discussions suggested other variables that should be included in the design for data collection. For example, the apparent difference between schools' reactions to decision situations externally mandated and those internally proposed suggested that the genesis of the prompt might be an important variable in our analysis. Similarly, we noted that the type of decision might affect the pattern of decision making. We were also attentive to the role that key individuals (such as the school principal) might play in the decision process.

Category systems were generated for classifying each of the three key variables -- the type of decision, the relevant personnel, and the kinds of information that were brought into play, and coding schemes were developed for other variables of interest -- the genesis of the prompt, the number of options, the length of the decision sequence, the existence of a strategy for decision making, and the identification of the issue under consideration with a particular group of people. After several drafts, the revised form was pilot tested and any remaining ambiguous items or confusing language were eliminated. (See Appendix E.)

On the forms the coders were asked to make two critical evaluative judgments about the interview itself. One concerned the level of missing data; the other reflected the accuracy of the sequencing of events.

From the written summaries we learned that not all the interviews explored the significant occurrences in equal detail. Some respondents were unable or unwilling to carefully reconstruct the school's activities

related to the significant occurrence they identified. Some interviewers devoted so much time to context-setting discussions or dialogues about the respondent's general experience with evaluation that the exchange related to each specific occurrence was extremely abbreviated. In order to assess the completeness of the descriptions, one item on the coding form asked the coders to judge the amount of information they believed to be missing from the portrayal.

The first data synthesis efforts also showed that the respondents did not always recount incidents in precise chronological order. The interviewers' follow-up probes often uncovered details that had to be inserted into the skeleton sequence of events which was emerging. While most interviews finally arrived at a clear ordering of events, (though it may have been derived in a jumbled fashion), in some instances the sequence of events was never clarified. Either the respondent could not remember the exact sequence or could not be guided into clarifying the order of events. Even when the respondent was cooperative the interviewer did not always recognize an inconsistency or lack of proper sequencing during the interview and failed to ask for clarification. Therefore, the coders also were asked to rate their confidence in their reconstruction of the sequence of events.

Ensuring the Reliability of the Coding Process. We instituted a number of procedures to ensure that the data would be coded reliably. Only four coders were used: each was a doctoral student in evaluation. The coders were involved in the development, revision, and pilot testing of the coding forms, thus insuring that any conflicting interpretations and confusions about language were clarified before the coding

began. Their complete participation helped to standardize the coding process.

Most importantly, the coders worked in pairs. As they listened, the two coders filled out a single coding form, replaying the tape when their interpretations differed to adjudicate their differences. The coding pairs were periodically rematched so that no pair-dependent interpretative biases entered the analysis. The coders could review the initial written summary first in order to listen for greater detail the first time the tape was reviewed. In addition, the same pair listened to all three tapes that had been made at a given school and listened to the tapes in the same order in which the interviews had taken place.

Finally we made one empirical check of the reliability of the coding process. The same set of tapes were reviewed by different pairs of coders and their results were compared. This comparison showed that the scoring was essentially the same. Small differences existed, but these were mostly in terms of degree. That is to say, one pair identified an input into the decision process as "classroom tests" while another identified it as "tests, undifferentiated". However, the sequencing of events and qualitative judgments about the accuracy of the descriptions were the same.[7]

Based on the precautions that were taken in developing the coding forms and procedures and the results of this *post-hoc* comparison, we felt secure that the refined data base reflected the descriptions that had been given by our respondents. We knew from our earlier external

[7] These differences in degree indicated to us that it was not possible to make the fine differentiations that were included in our category systems, and in the final data analysis we grouped responses at a higher level of aggregation.

validation that the respondents believed the content of the interviews accurately reflected the events that had transpired at their schools.

Quantitative Analysis

Overview. The variables we selected for quantitative analysis and the relationships we chose to investigate were in large part derived from our initial qualitative analysis -- the written summaries, the group discussions, the Human Data Bank, etc. Though we hoped to gain new insight from the numerical comparisons, our guiding principle throughout was not to sacrifice descriptive accuracy in the name of quantitative efficiency.

The analysis proceeded in stages. First, we classified the significant occurrences into categories that reflected the subject or action under consideration. The individual decision sequences previously had been coded in terms of the personnel involved and the type of information used at each step, so we then developed categorizations for the variables "personnel configuration" and "type of information". Finally, we analyzed the relationships between the three variables. Using 'type of information' in the role of dependent variable we examined the decision sequences to see if there were any identifiable relationships between the information profiles and the type of occurrence or the configuration of personnel involved.

The evolution of the significant occurrence had been conceptualized in four chronological phases -- recognition of a prompt to action, decision making, ratification of the decision and dissemination. Most of the interaction occurred in the decision making phase, and our analysis was concentrated there.

The subsample. Our written summaries contained descriptions of 109 significant occurrences.[8] A number of factors intervened to reduce the size of our final data base. First, "significant occurrence" was not synonymous with "school decision". In some cases our respondents perceived their school to be so dominated by external factors (e.g., district-wide integration requirements) that they only identified significant occurrences in which the school essentially had no options -- their only choice had been to comply with the rules. Our interviewer probed to determine if there were other events the respondent judged to be significant, events in which the school had some latitude for action. In 20 cases we were not able to elicit two such occurrences. As a result we did not always obtain two significant occurrences in which there had been some within-school choice of action.

Second, the focus of our interviews had been on factors that affected evaluation utilization in the context of each occurrence. This investigation was usually accomplished by reconstructing the sequence of events that had transpired. However, not all interviews proceeded in this manner. We knew when we decided to reconstruct decision sequences that not all of our descriptions would be complete in this regard. Thus we were careful to include a measure of the accuracy and completeness of the portrayal in our coding forms.

---

[8] We conducted 65 interviews, and hoped to obtain descriptions of two significant occurrences from each respondent. However, digressions, elaborations, time constraints and the inability of some respondents to identify any significant occurrences made our actual sample somewhat smaller.

The first step in the analysis was to identify a subsample of occurrences that contained complete descriptions of the school's decision making process. This subsample included 73 school decisions, and these 73 cases formed the basis for all the analyses which are reported in the following chapters.

Analytic procedures. The quantitative analytic procedures we employed for investigating the data were not complex; most of our analysis consisted of frequency counts and cross tabulations. There were three reasons for this. First, we were not looking for obscure relationships that would be difficult to detect. Our research questions were directed toward naturally occurring patterns among pairs of variables. Second, we did not have interval or ratio scales that could be subjected to more sophisticated statistical analyses; our data were categorical -- different types of information, personnel or occurrences. Third, though we began with 65 interviews, when we separated them into natural categories, the number in each cell of the analysis was too small for most statistical techniques. Consequently most of our analysis consisted of frequency counts and cross tabulations.

The advantage of this type of analysis is that the results are very easy to understand. We sorted the decisions by type of occurrence and compared the different information profiles that were observed in each. Similarly we sorted by type of personnel and compared information use patterns. In addition, we cross-tabulated information use against the other variables of interest we had coded -- source of the prompt, number of options, length of the decision sequence, existence of a strategy for decision making and identification of the issue with a particular group.

One disadvantage of the small number of observations in many of the categories of our analysis is that it prevented us from conducting any tests of statistical significance on the differences in information use we observed. This was less a drawback than one might imagine, however. One must not forget that all the quantitative analyses were derived from our original qualitative inquiry, and we already had some insight into which effects were significant from our extensive earlier review of the data. Newly discovered quantitative differences would have to correspond with these prior understandings before we considered them to be reliable.

However, we had to develop some guidelines for judging the importance of the differences we might detect. We established the following guidelines: (1) Place little emphasis on differences that were detected when the number of cases under consideration was under five; they had limited reliability. (2) Use the average information use profile across all cases in the sample as the baseline for testing each category. Put little emphasis on differences that are less the magnitude of this overall average. (3) In all instances use earlier insights and the knowledge gained from the initial data synthesis as the final arbiter of the importance of differences that were detected.

Chapter 3

DESCRIPTIVE STATISTICS: DECISION MAKING PHASE

INTRODUCTION

In Chapters 3 and 4 we will focus on the decision making phase, and present the results of our analysis of this segment of the school's activities. We emphasize the decision making phase because most of the information conveyed by our respondents related to the decision process. The analysis of the recognition, ratification and dissemination phases will be included in Chapter 5.

The results presented in this chapter are primarily descriptive, focussing in turn on the three variables, type of occurrence, type of information, and type of personnel. Each section presents the results and also includes some preliminary discussion of the meaning of the data. We hope, thereby, to avoid the "symbol shock" which can follow lengthy presentation of numbers and figures. Nonetheless we reserve our overall comment and conclusions for the final chapter.

BREAKDOWN OF SIGNIFICANT OCCURRENCES

The respondents in our sample were asked to identify significant occurrences for discussion. Much can be learned from that identification about the local school decision maker's perspective on important school events, the scope of program change that commonly occurs and the kinds of activities on which evaluation might conceivably be brought to bear.

We reviewed the list of significant occurrences carefully and classified them into general categories. Some occurrences involve aspects of more than one category. For example, planning a new Title IVc program involves considerations of both budget and the distribution of administrative staff time. However, we tried to assign each occurrence to the category that most appropriately reflected the primary thrust of the activities described.

At this point analytic efficiency argued for constructing four or five broad topic areas that would subdivide the sample more or less equally. Unfortunately, the naturally-occurring similarities among the occurrences did not create such a breakdown. There were a dozen identifiable clusters of decisions ranging from purely administrative, such as hiring new staff members, to the instructional, such as developing a special classroom arrangement for students who fall behind in their reading program. The complete list included occurrences related to: (1) instrumental materials, (2) creation of new programs, (3) out-of-classroom professional staff, (4) small scale instructional programs, (5) bilingual program implementation, (6) general curriculum guidelines, (7) miscellaneous activities, (8) personnel actions, (9) evaluative events, (10) parent involvement, (11) staff development, and (12) patterns of student grouping for instruction. Though the size of these groups varied greatly and some were so small as to preclude reliable tests of differences in later analyses, faithfulness to the situation we were trying to depict required that we maintain all 12 categories. Full descriptions of the 12 categories are as follows:

1. *Instructional Materials* (INS MATL)[9]

As opposed to curriculum guidelines, significant occurrences grouped under the heading of instructional materials did affect classroom instruction directly.

> Example: After many unhappy years with the DRP program -- a phonics-based developmental program, the school decided to purchase a new basic reader for the following year (19SP1).

Each of these occurrences related to supporting instructional material that the teachers used on a daily basis. This category does not include any actions to change teachers' pedagogical styles directly.

2. *Creation of New Programs* (NEW PROG)

This category includes all instances in which additional funds or staff time was available for development and implementation of a new instructional program.

> Example: A new program was instituted in selected schools in Metro District this year. It was designed to provide extra instructional activity at the conclusion of a regular school day for students who were in heavily racially-isolated schools. Teachers were given an eleven percent salary bonus and asked to provide seven additional hours of student contact per week. While some possible forms for this after-school activity was suggested, each school could determine on its own the type of program it would provide. (04P)

3. *Out-of-Classroom Professional Staff* (STF PERS)

This group of occurrences involved changes in the roles and responsibilities of out-of-classroom staff. Other instances in this category included changing or expanding the role of other auxiliary staff positions, such as school psychologists or a multi-cul-

---

[9] The phrases in parenthesis represent the abbreviated eight-character labels that were retained by the computer and used in the charts and figures that are reproduced later in the chapter.

tural education coordinator.

> Example: In the past the schools' four resource teachers, who were subject matter specialists, worked only with selected students on a pull-out basis. To lower the pupil-to-teacher ratio for reading instruction, all the resource teachers were assigned to work with a regular full sized group of students every day during the reading time period. (04SP1)

4. *Small-scale Instructional Program* (SML INST)

This category includes occurrences that affected instruction in only a small number of classrooms or only a small number of selected students.

> Example: This school had a half-day pre-kindergarten program. The school decided to allow one of the teachers to work out a reading readiness program for the pre-kindergarten and first grade students and eventually adopt it.

This category also includes occurrences that affect the whole school, but only in a minor way. Changes in the once-a-month multicultural program exemplify this latter group.

5. *Bilingual Program Implementation* (BILINGL)

This category includes those occurrences that related to the implementation or expansion of bilingual programs.

> Example: The number of Hispanic students enrolled in this school had been increasing slowly over the past two or three years. Bilingual aides had been used to help with the language needs of those few students who could not communicate effectively in English. As the number of LES/NES students increased, the school could no longer provide effective instruction using only aides, and they decided to adjust their staff allocation so a full-time bilingual teacher could be employed to work with those students who needed a bilingual program. (17SP2)

The category also includes occurrences relating to the provision of bilingual instructional materials as well as occurrences focused on increases in the number of bilingual staff.

6. *General Curriculum Guidelines* (GEN CURR)

These decisions involved changes in the official curriculum or general guidelines for instruction.

> Example: In the past there had been separate curriculum strands for monolingual and bilingual students. The school decided to unify the two strands into a single curriculum and adopt common grade level objectives for all students (02 P1).

These general curriculum decisions do not represent any attempt *directly* to supervise day-to-day instruction or to alter the teachers' pedagogical approaches to students. They deal with purposes and goals rather than means or methods to achieve them. They are distinguished from occurrences involving selection of new instructional materials or changes in program guidelines that affect only a small segment of the school.

7. *Miscellaneous* (MISC)

Miscellaneous occurrences include a variety of activities of lesser instructional importance which did not fit under any of the other categories. Included in this category are occurrences involving changes in scheduling of auxiliary school activities, such as festivals or dismissal times, decisions about the timing of a mandated activity, or decisions about clerical or paraprofessional staff.

8. *Personnel Actions* (PERSONNL)

This category includes those occurrences that were primarily related to the principal's administrative role in hiring, firing, promoting or transferring personnel.

> Example: This school qualified for Title I funding for the first time the previous year. The principal had to select one person on the staff to serve in the newly

created administrative position of Title I Coordinator. (26SP1)

9. *Evaluative Events* (EVAL)

The small number of significant occurrences in this category related to testing or the needs assessment process.

> Example: The regular tests that accompanied this school's phonics-based reading program included a number of nonsense words, and many of the teachers objected to using these in measuring the student's achievement. The teachers omitted such words from their instructional program and felt they were inappropriate. After some discussion the school decided to eliminate nonsense words from the tests and adjust the scoring system accordingly. (15SP1)

10. *Parent Involvement* (PAR INVL)

In this category the significant occurrences involved activities directed toward greater participation or communication with parents.

> Example: In the past this school has offered workshops for parents in a variety of subjects. Attendance has been low and they have only had limited success. The school decided to make modifications in the parent-training program in order to improve its effectiveness. (03SP2)

This category does not include activities related to the responsibilities of classroom aides, a paraprofessional staff position frequently filled by parents.

11. *Staff Development* (STFF DEV)

This category includes those significant occurrences which involved improving the professional qualifications of the staff.

> Example: This school had a sufficient number of bilingual teachers to meet its legal commitment to LES/NES students. However, many of the monolingual teachers wanted to be able to communicate better with the Spanish-speaking children in their rooms. As a result, they organized a voluntary after school Spanish class for faculty under the auspices of the staff development program. (02SP2)

Orientation to new program guidelines as well as special training are included under this heading.

12. *Patterns of Student Grouping for Instruction* (STU GRPS)

This category included those significant occurrences that centered on the instructional grouping of students.

> Example: Last year the second grade teachers at this school reorganized their instructional program into a 'departmentalized' structure, in which each taught particular parts of the curriculum to all students at the grade level. After reviewing low test scores they decided to return to self-contained classrooms.

The occurrences of this type were almost evenly divided between instances in which instruction had been carried out in self-contained classrooms and was subsequently transformed into another arrangement -- team teaching or departmentalization -- and those instances in which the change had taken place in reverse order.

The breakdown of significant occurrences by categories is shown in Table 1 . The frequency of each type of occurrence is displayed as well as the percentage of the total sample that fall into each category.

TABLE 1

Type of Significant Occurrence

CATEGORY		NUMBER OF OCCURRENCES	PERCENT OF TOTAL
INS MATL	(Instructional materials)	13	17.8
NEW PROG	(Creation of new Programs)	11	15.1
STF PERS	(Out-of-classroom professional staff)	8	11.0
SML INST	(Small-scale instructional program)	8	11.0
BILINGL	(Bilingual program implementation)	7	9.6
GEN CURR	(General Curriculum guidelines)	5	6.8
MISC	(Miscellaneous occurrences)	5	6.8
PERSONNL	(Personnel actions)	4	5.5
EVAL	(Evaluation-related occurrences)	3	4.1
PAR INVL	(Parent involvement)	3	4.1
STFF DEV	(Staff development)	3	4.1
STU GRPS	(Patterns of student grouping for instruction)	3	4.1

Discussion

The information summarized in Table 1 elicits a number of observations:

1. The vast majority (64%) of the significant occurrences identified by our respondents concerned matters of curriculum and instruction (GEN CURR, INS MATL, STU GRPS, NEW PROG, BILINGL, SML INST). In this regard they share what would be considered the common view of what is "important" in schooling. These are also areas in which evaluation can conceivably have positive impact.

2. On the other hand, there were a fair number of respondents who identified non-instructional actions as significant. Seventeen percent of the occurrences involved personnel actions, parent involvement, and other miscellaneous occurrences of limited instructional significance. Some of these 'significant occurrences' were rather trivial in nature.

   The fact of the matter is that some of the administrators we talked with focused much of their attention on relatively small aspects of their jobs. This group included some principals who were "coasting" toward retirement and focused on minor administrative matters rather than large-scale program innovations. But it also included, for example, some resource teachers who had limited areas of responsibility and consequently narrower views of school decisions.

   Furthermore, a few active decision makers proffered very unimportant activities as "significant occurrences". One reasonable explanation for this may be captured in Weiss's (1980) observation that decisions are not *made* at schools but rather "accrete" indirectly over time. Thus individuals may not identify any major decisions and not feel any school actions were significant. A sense of "impotence" may also account for the identification of unimportant activities as significant occurrences.

3. In fact, much of what occurs in the schools *is* prompted by forces outside the control of the individual school administrator. Forty-five percent of the significant occurrences that were described to us had their genesis in external events. Changing

demographic patterns were a chief source of activity in the schools. Similarly, the legal maneuverings surrounding school desegregation and the district's actions in this issue had strong impact on the individual schools. As one staff person told us,

> I think right now the judge is effecting as much change in education as anyone. The law dictates. Decisions are made that schools are asked to live with that they may not be capable of dealing with effectively. Yet we're asked to more and more. (12SP2)

Thirty-five of the original 109 significant occurrences that were identified by our respondents related to changes in the district's integration and bilingual programs. For example, most of the instances in which new programs were initiated (Category 4) involved a district-level attempt to provide additional assistance to schools that could not be desegregated by pupil transportation.

Understandably, some of our respondents felt that much of what was done at the local level was prescribed by program regulations and the district administration. Some of the decision makers in our sample viewed their own role as purely reactive. The following comments typify this perspective:

> It just seems we've been bogged down doing the mandates of decisions made higher than the local schools. Certainly our last couple of years have been spent adjusting to new mandates, new laws that have just been thrust upon us. (14P)

> The coordinator from the Area Office hands it down and, of course, we go along with it. It was not something we could decide ourselves. When they say go, we go. (17SP1)

All in all, it is fair to say that much of what occurred in the schools during the year in which we conducted our interviews

involved school responses to external events, and prominent among these were district directives.

4. One can, however, react to such events in entirely different ways. Some principals felt overwhelmed; others did not. The latter group saw one of the main tasks of their job as figuring out ways to accomplish what they wanted despite the flood of regulations. Sometimes external mandates were even helpful -- they gave the the administrator extra weapons in his or her desire to bring about change. This suggests that 'externality' per se does not imply limitation. Rather the level of opportunity for action seems to be a function of how one perceives the situation and chooses to respond.

5. While the overwhelming majority of the significant occurrences had to do with elements of the instructional program, none involved direct attempts to influence the manner in which individual teachers carried out instruction. There were changes in guidelines, management systems, text books and diagnostic tests, but there were no clear instances in which the professional boundary separating administrative functions from instructional decisions was broached. The classroom door, for all intents and purposes, remained closed.

This observation should not come as a surprise, and we offer it only as further description of our sample. A currently popular theory describes schools as "loosely coupled" systems in which there is marked separation between the administrative sector and the sector that actually delivers the services (Weick,

1976; Meyer & Rowan, 1977). Similarly Miles (1980) suggests that there are different "zones" of decision making within schools, and that instructional decisions fall within the teacher's zone. Our sample of significant occurrences tends to add credence to these theoretical descriptions. There were no clear instances in which the zone of instructional decisions was open to direct action from the administration.

6. It is particularly interesting to examine those significant occurrences that related to the development of new programs. They are a special sub-sample because they represent instances in which the normal constraints on action have been relaxed. Teachers and principals usually report that their options are limited by myriad pressures: scheduling constraints, budget constraints, rules and regulations, and the like. In most of the instances in this category, the school had wide latitude to innovate as this description shows,

> Did the teachers have any constraints in deciding what type of program they would like to initiate...? Very little. Each teacher could have their own written proposal which was submitted to the principal for approval so they had a great deal of freedom...(07SP2)

Yet the amount of innovation was almost nil. Typically the additional hours that were required of teachers in the racially-isolated schools were given over to small-group tutoring or to special-interest clubs. While we are not suggesting that either of these two activities is inappropriate, it is interesting to note there were no instances in which standard instructional patterns were abandoned for something unusual, creative or daring.

7. Finally, we note how few of the decision makers viewed evaluative events as significant. Evaluation per se is simply not a matter of great significance in the schools. Daillak's (1980) research in Metro indicated that the impact of evaluation was limited, and we did not expect that many of the decision makers would identify evaluation-related occurrences among the most significant activities that had transpired during the previous year.

BREAKDOWN OF TYPES OF INFORMATION

We used the order in which events occurred to organize our reconstruction of the school's decision process. At each identifiable step in the decision sequence we asked for information about two components -- the personnel who were actively involved and the information sources that were brought to bear on the interchange.

Ideally, a respondent might describe a meeting in which certain informed individuals discussed data from different sources in order to illuminate a question and select the best course of action. In such a situation one could define the notion of "information" very narrowly as facts derived from direct observation of a relevant situation, scientific analysis of many situations -- i.e., research and evaluation -- or from collegial reporting of similar situations. In reality, however, much of what transpired in such meetings was not merely an exchange of distinct facts, but rather an exchange that also included personal opinions, attitudes and beliefs. These opinions were no doubt in some manner derived from direct experience, scientific analysis, contact with

others, and the like, their exact genesis was unknown. Our respondents often were unable to analyze the process through which these beliefs and opinions were formed. The comments of one principal allude to this distinction,

> Observation and visiting classrooms, labs, feedback from teachers and I guess this gets down to an individual thing, teachers expressing frustration or concerns about individuals working with their children, and then through my own observations, that helped make the decision. (10P)

As a result we expanded the definition of information, to include beliefs and opinions as well as pieces of data. We use the term 'type of information' to refer to the smallest descriptive units we could obtain relevant to the interaction. We distinguished and coded 28 types of information which were then grouped into 11 categories that contain inputs of a similar nature. The 11 categories are: (1) beliefs and opinions, (2) program requirements and budgets, (3) direct observation, (4) parent input, (5) district staff, (6) needs assessment, (7) external consultants, (8) tests, (9) collegial advice, (10) other evaluation activities, and (11) other information types. Full descriptions of these categories are as follows:

1. *Beliefs and Opinions* (OPINION) This category includes those instances in which the personal opinions or beliefs of a teacher, principal, or staff person were cited as important factors at a particular step in the decision process. Nothing further was known about the genesis of that belief or opinion.
2. *Program Requirements and Budgets* (PROG REQ) References to guidelines or regulations governing a program that were a factor at some point in the decision are included in this category. In

some instances, rules governing allowable expenditures or expenditure limits entered into the decision process, and they are included in this category as well.

3. *Direct Observation* (OBSERVTN) This category includes references to an individual's first-hand observations which were reported as evidence on a particular issue.
4. *Parent Input* (PARNT IN) This category includes input from parents, whether it came from representative parent committees or through informal contacts with staff.
5. *District Staff* (DIST STF) Advice and direction from Metro district staff constitutes this category. This includes the subject matter specialists as well as individuals in the administrative hierarchy, but it excludes people from the Evaluation and Testing office.
6. *Needs Assessment* (ND ASSMT)[18] This category includes instances in which information collected as part of a needs assessment was referred to in a particular decision. Most schools conduct a single, annual needs assessment to meet state program guidelines. Some schools conducted smaller-scale needs assessments at other times and these are also included in this collection of information.
7. *External Consultants* (EXT CONS) In some cases the schools requested information or advice from outside consultants and specialists. These inputs are included in this category. Publishers' representatives are also included in this category

[18] Evaluative inputs have been subdivided into three categories because evaluation is of particular interest.

along with other experts consulted by the school staff.

8. *Tests* (TESTS) This category includes all referemce to test scores. It includes the required, annual Title I achievement tests, classroom tests and other miscellaneous testing that respondents mentioned.
9. *Colleagial Advice* (COLLEAGS) Professional colleagues often exchange information, and this was cited as a factor in some decisions. References to information from principals or teachers at other schools is included in this category.
10. *Other Evaluation Activities* (OTH EVAL) The largest number of entries in this category referred to input from the evaluation consultant from the Metro E & T office. Additionally references to local evaluations and references to the results of the state PQR team review are included in this category.
11. *Other Information Sources* (OTHER) The category includes those few information inputs which could not be classified into any of the other ten categories.

The breakdown of information type by category is shown in Table 2 The frequency of each type of information is displayed as well as the percentage of the total sample that falls into each category.

TABLE 2

Information Sources

CATEGORY		NUMBER OF OCCURRENCES	PERCENT OF TOTAL
OPINION	(Beliefs and Opinions)	234	50.0
PROG REQ	(Program Requirements and budgets)	54	11.5
OBSERVTN	(Direct Observation)	39	8.3
PARNT IN	(Parent Input)	30	6.4
DIST STF	(District Staff)	27	5.8
ND ASSMT	(Needs Assessment)	26	5.6
EXT CONS	(External Consultants)	24	5.1
TESTS	(Tests)	13	2.8
COLLEAGS	(Collegial Advice)	11	2.4
OTH EVAL	(Other Evaluation Activities)	9	1.9
OTHER	(Other)	1	.2

Discussion

Table 2 illustrates some interesting relationships that are worthy of further comment:

1. Far and away the largest single input into decisions was beliefs and opinions. This can be interpreted in a number of ways. It might simply reflect the respondents' lack of knowledge and insight about the reasoning process of others. Another interpretation would argue that peoples' core values and attitudes form over extended periods of time as a result of a multiplicity of experiences and consequently do not have identifiable short-term causes.

A third perspective is provided by Lortie (1975), who portrayed teaching as a particularly isolated profession that had an insulated, cellular quality. Teachers are expected to learn how to teach from their own personal experiences without relying on input from others. Thus, personal experience and personal opinion become elevated in importance. One could easily argue that the natural extension of this pattern of socialization to the profession is a lowering of the interest in and reliance on exchanges of facts and pieces of data between teachers and an increased emphasis on the importance of self-derived attitudes and opinions. Lortie's perspective is echoed in these remarks,

> I guess the most important thing is my experience as an educator. I think that we do not have a body of experimental knowledge that we can call on and say "this is clear cut". So I think in terms of looking at the school day and such kinds of things we do with children...I really don't have anything to base it on. My experiences as an educator.... (19P)

What we observed in this study is probably a combination of all these forces. Whatever the case, we can see clearly that beliefs and opinions are important. We will consider the role that evaluation might have in opinion formation in a later discussion.

2. Frequent citation of program guidelines and regulations adds weight to some principals' contention that their hands are often tied. A number of administrators in our sample felt they operated in a universe of limited options.

> That's right...in many cases it's a joke to say that there are choices...the choices you have are not significant enough to make any difference...They would be better off not telling us we have a choice when in fact we don't. (12SP2)

The importance our respondents gavee to rules and guidelines in their accounts of the significant occurrences tends to corroborate that point of view.

3. The paucity of tests and other evaluation inputs is discouraging but not surprising. There were very few instances in which tests or other evaluations were cited in these significant occurrences. It seems that little has changed in this area since research on the subject of evaluation utilization began in the early 1970's.

4. On the other hand, needs assessment data were brought to bear on an important school decision twice as frequently as tests. This adds some credence to the belief that needs assessment can have a key role in school planning (even if its initial use is forced upon the school). Here is a case in which it was useful,

> Budget cuts necessitated making other changes, according to the principal so he gave them (the staff and parents) a needs assessment. He has discussed the needs assessment process and one of the needs that was being assessed...the staff and parents decided that we didn't need a reading coordinator or math coordinator, that they would rather see people in classrooms working with individual children. So we eliminated both positions. (10P)

PERSONNEL CONFIGURATIONS IN SCHOOL-LEVEL DECISION MAKING

A wide variety of personnel was involved in the various school decisions we investigated. Initially our coding form listed 20 different personnel groupings, but as we listened to the tapes this list grew to more than 30 different configurations of personnel described by our respondents.

We tried to manage this diversity by matching similar configurations. We identified seven categories: (1) administrators, (2) whole staff, (3) teacher-administrator groups, (4) teachers (5) parent-aide-staff groups, (6) parent-aide groups, and (7) consultants. We classified each personnel group that was reported to us into one category as follows:

1. *Administrators* (ADMIN) This category consists of instances in which either the principal or various "staff persons" were involved at a particular step in a decision. The size of the administrative group does not matter.
2. *Whole Staff* (STAFF) This category includes those instances in which the whole staff met as a group at some point in the decision making process. We made no distinction between issues that were included on a planned agenda and discussions that occurred spontaneously in staff meetings.
3. *Teacher-Administrator Groups* (T+AD GPS) An executive or leadership committee is an example of a teacher-administrator group, one that is formally constituted and has official status at the school. In addition, this category also includes informal groups of teachers and administrators and informal groups "dominated" by teachers and administrators. That is, we have included in this category one or two instances in which an informal group of teachers and staff persons also included a small number of classroom aides, clerical personnel, or parents. If the group was clearly dominated by the school professionals, it was included in this category.

4. *Teachers* (TEACH) This category consists of instances in which either individual teachers or groups of teachers were cited as being the personnel involved at a particular stage in the decision process. We include citations for individual teachers, citations for informal groups of teachers, and the citations for representative teacher committees.

5. *Parent-Aide-Staff Groups* (PAR+STFF) Parents and aides were usually brought into the decision process in mixed groups with school staff. The school site council parent-teacher conferences are examples of such groups. In contrast to T+AD GPS with some parent participation, this category includes groups in which parents played the sole or predominant role.

6. *Parent-Aide Groups* (PAR+AIDE) This category includes instances in which parents and/or classroom aides participated singly or in groups in the decision process. This includes individual parents, individual aides, formal parent committees, informal parent committees, and instances in which the total parent population was surveyed about their opinion. We included such a diverse collection in this category because the number of instances in which any of these parent or aide configurations were cited in the decision process was very small.

7. *Consultants* (CONSULTS) Under the general heading of consultants we include administrative staff from the downtown office, evaluation consultants from the Evaluation and Testing Office, district subject matter consultants, and external consultants selected by the school. (Representatives from instructional

materials companies and book publishers comprise most of the latter group.) This category includes consultants described as acting singly in the decision process, and the consultants who met with groups of staff persons or teachers.

The frequency with which each different personnel category entered the decision process is shown in Table 3 The percentage of the total number of citations that belong to each category is also presented.[11]

TABLE 3

Personnel Configurations

CONFIGURATION		NUMBER OF OCCURRENCES	PERCENT OF TOTAL
ADMIN	(Administrators)	63	30.0
STAFF	(Whole staff)	62	29.5
T+AD GPS	(Teacher-administrator groups)	36	17.1
TEACH	(Teachers)	25	11.9
PAR+STFF	(Parent-aide-staff groups)	16	7.6
PAR+AIDE	(Parent-aide groups)	5	2.4
CONSULTS	(Consultants)	3	1.5

[11] It was possible for a personnel group to enter a decision more than once; in compiling Table 3 we counted each of these steps separately. For example, if a matter was discussed at three different meetings of the full school staff, this would be counted three times under the category of staff. In reality there were few cases in which a personnel group entered a decision more than once; multiple entries occurred in less than one-quarter of the personnel citations. Thus the relative balance exhibited in Table 3 is not strongly biased by a few multiple instances.

Discussion

1. Of particular interest to this study is the very limited number of times that district consultants participated directly in decisions. Subdividing the consultant category into its component parts, we discovered no instances in which personnel from the Evaluation and Testing office participated directly in the decision process. Area staff were mentioned occasionally, as were subject matter consultants, but members of the E & T unit were not directly involved in any of the decisions described to us.
2. Similarly, it is interesting to note how seldom parents and classroom aides are cited as being *directly* involved in the decision process. Yet we know from Table 2 that their ideas were incorporated indirectly. This suggests that the parents' role in the formal decision mechanism is small, but that their ideas are informally communicated to members of the staff and do get considered when program decisions are made.
3. The bulk of the decision making involved the active participation of the whole professional staff. There was an overall balance between administrators and classroom teachers. In fact, there were very few instances in which decisions were made solely by administrators or solely by teachers.

Chapter 4

INTERRELATIONSHIPS: DECISON MAKING PHASE

INTRODUCTION

In the preceding chapter we looked at univariate analyses of the three key variables: type of decision, type of information, and personnel configuration. In this chapter we present the results of three bivariate analyses: the relationships between type of information and type of decision, between type of information and personnel configurations, and between different types of information. In addition, the relationships between type of information and the synthesis (number of options, length of decision, strategy for decision making, genesis of prompt, and group identification) are included.

Because the data are categorical we could not compute correlation coefficients; rather, we examined graphical displays of cross-tabulations between the variables. We also compared the pattern of information use on each individual variable with the pattern of information use in the total sample.

One word of explanation seems in order before we proceed. The bar graphs which present the data in this and subsequent sections are scaled differently from the tables that were used previously. Tables 1, 2, and 3 showed absolute frequencies and percentages. In contrast to this, the bar graphs which follow are based on the mean number of occurrences of each category of information *per decision*. This normal-

izes the displays and makes it possible to compare the information profiles. Unfortunately, the computer graphics program that was used to generate the figures could not accommodate mean values less than one. As a result the data were multiplied by 1,000 so that they no longer appeared as decimals. Thus the figures themselves will display the *mean incidence* of a particular information type that would occur if there had been *1,000 decision sequences*. The average information use in a single decision can be obtained by dividing by 1,000.

DECISION TYPE VS INFORMATION TYPE

Figure 1 displays the average level of information use for the entire sample of 73 significant occurrences. The relationships between the various information categories are exactly the same as those portrayed in Table 2; only presentation and scales differ.

Figure 1 will serve as the baseline against which all the other information profiles will be compared. We examined each of the 12 decision types in turn and compared them with this baseline profile. In this section we will discuss only those instances in which the information profile differed significantly from the baseline profile. We were guided in this decision of significance by the principles that were outlined previously: not placing too much confidence in differences that are based on a very limited number of observations, nor on differences of lesser magnitude than the baseline value itself. First we will examine the decisions in which we found increased evaluation use, then those in which there was a significant decrease in evaluation use. Finally, we will discuss categories in which there was notable change in information

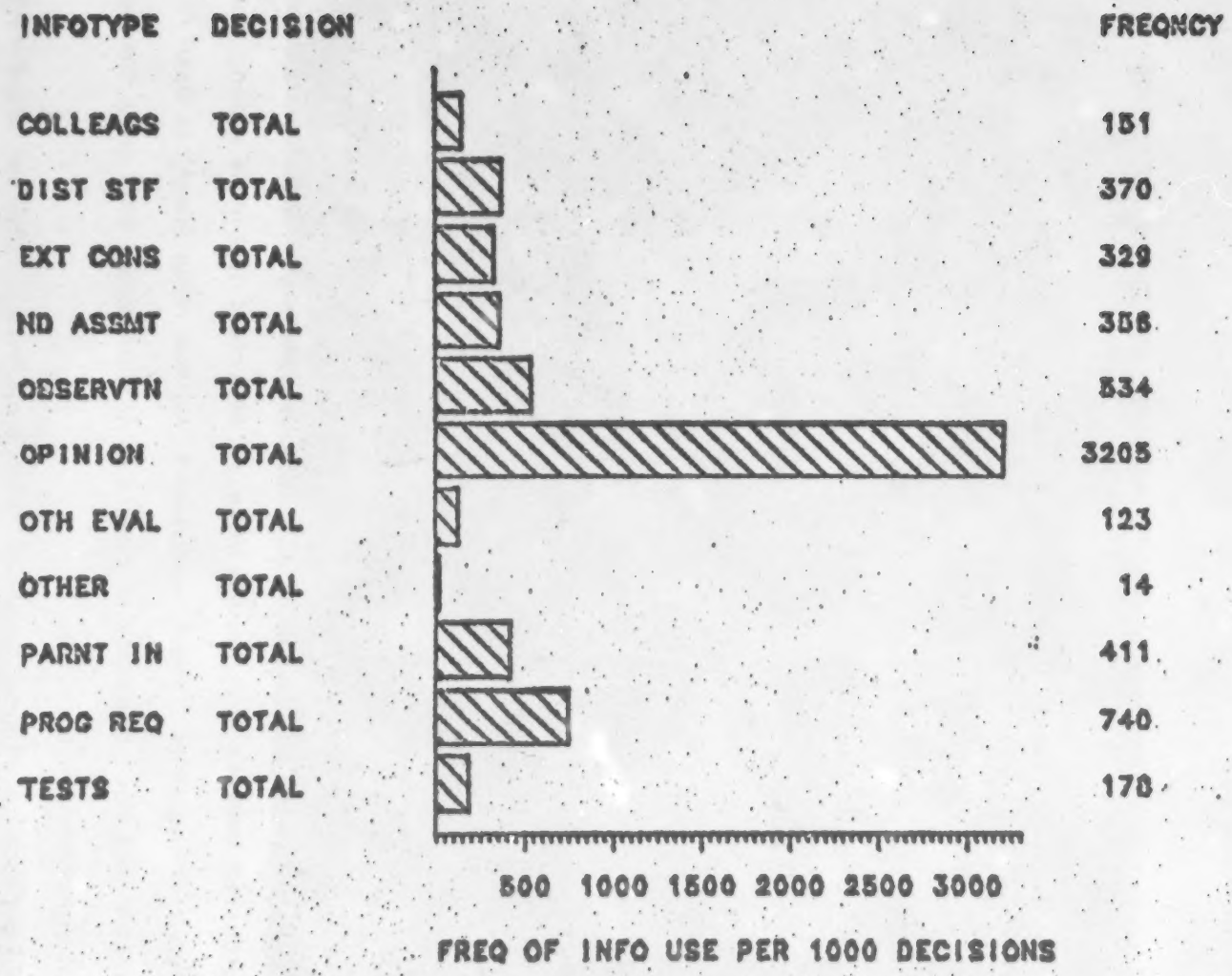

Figure 1: Information Use, ALL DECISIONS, N=73

use in areas other than evaluation. (All 12 comparisons will be found in Appendix F.)

Incidences of Increased Evaluation Use.

There were only two categories of decisions in which the incidence of evaluation use was markedly greater than the overall mean. These were decisions relating to general curriculum guidelines (GEN CURR) and decisions involving bilingual program implementation (BILINGL). (Increased incidence of evaluation also occurred in decisions involving student grouping, but the differences were not marked and the size of the sample was small.)

GEN CURR. In Figure 2 the pattern of information use for general curricular decisions is displayed alongside the baseline profile. In these decisions we observed a much higher than average reference to testing and to needs assessment. Looking more closely at the cases in the GEN CURR category, we found decisions to revise the number of reading levels that a student was supposed to accomplish in a grade level, and to unify a curriculum that had been split into distinct monolingual and bilingual strands. The increased references to tests occurred because test results were an important factor in making both types of decisions. Needs assessment, on the other hand, is itself the factor that caused people to recognize the problems that were the foci of significant occurrences.

BILINGL. Figure 3 shows the comparison between occurrences involving bilingual programs and the total sample. A similar pattern of increased reference to needs assessment was found among those signifi-

cant occurrences relating to the school's bilingual program. Typically, in these instances, the awareness that there was a need for a change came as a result of a language proficiency survey conducted as part of the school's needs assessment process. Recognition of demographic changes first crystallized in these annual needs assessments.

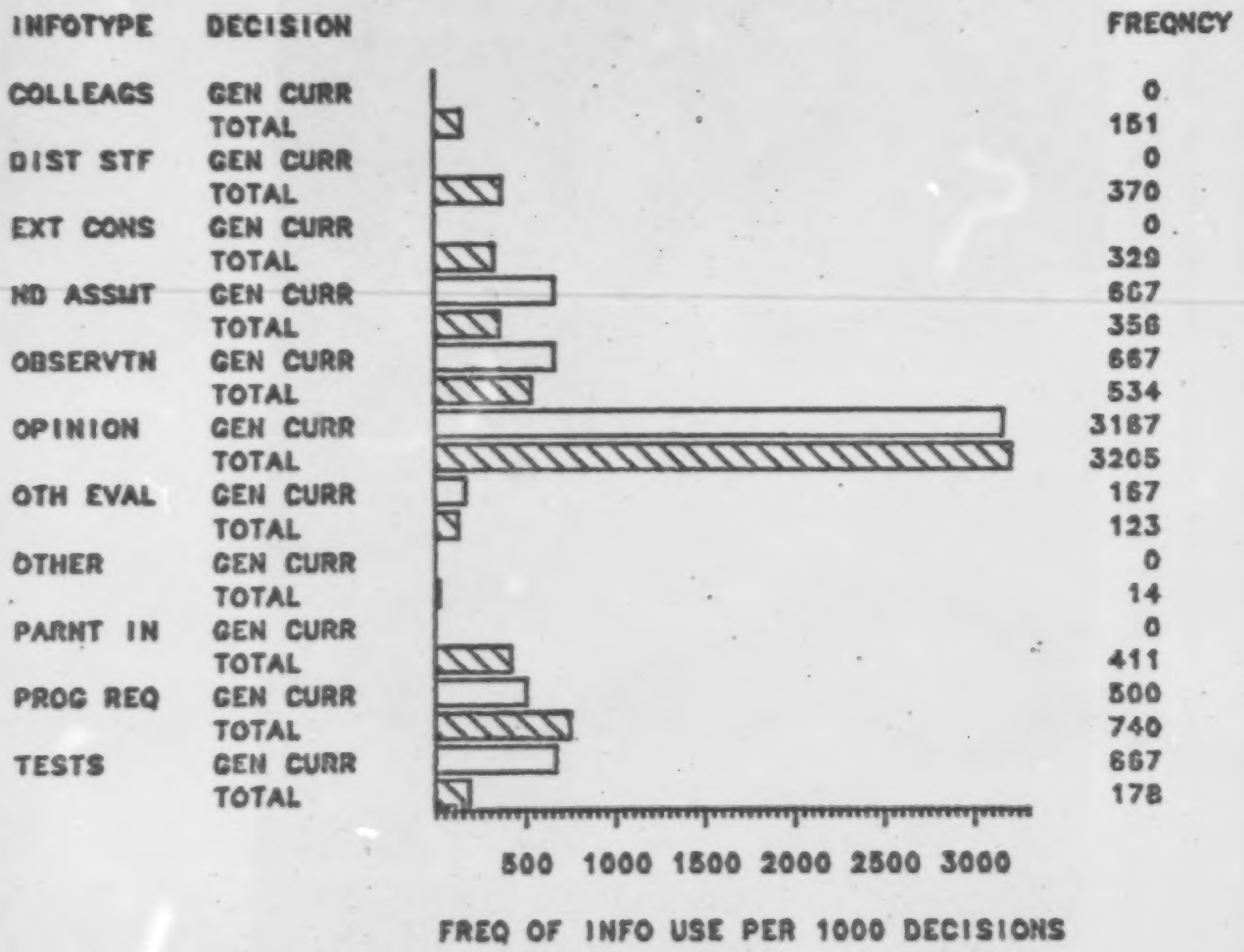

Figure 2: Information Use, GEN CURR (N=6) vs. TOTAL (N=73)

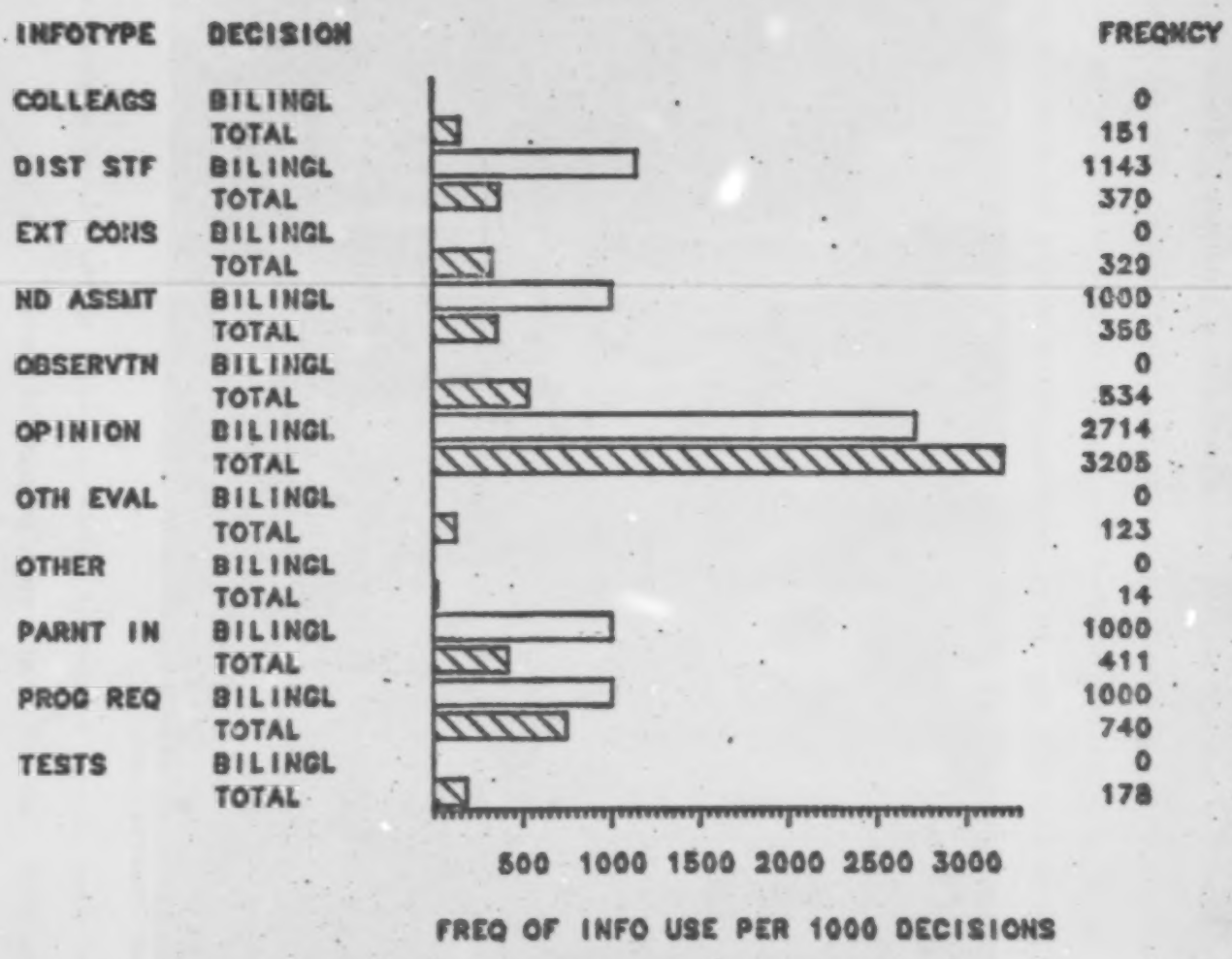

Figure 3: Information Use, BILINGL (N=7) vs. TOTAL (N=73)

Incidences of Decreased Evaluation Use.

The only case in which the evaluative categories were significantly lower than the average was in the miscellaneous category. There was no common thread among the five significant occurrences that fall into this category and comparison of the baseline profile of information use and the profile that applied to miscellaneous decisions failed to reveal any new insight.

AGGREGATE GROUPS. We should mention three other categories in which the use of evaluation was much lower than the average. Decisions relating to parent involvement, personnel actions, and staff development all displayed profiles in which the three evaluative categories fell well below the baseline. (See Appendix F.) If we aggregate these three categories into a single unit, it would have an acceptable sample size, and we could be comfortable drawing some tentative inferences. This aggregation is reasonable because PARNT IN, PERSONNL and STFF DEV all consist of non-instructional decisions. They deal with administration, supervision, and professional advancement, rather than classroom management, student performance, or instruction. It seems reasonable that decisions in non-instructional areas would seldom refer to needs assessments, tests, or evaluation of other types.

Other Observations.

Some other strong differences relate indirectly to the use of evaluative information. We will briefly mention some of these.

INS MATL. While the level of evaluative information that we find in decisions relating to instructional material (INS MATL) is about the

same as the overall level, there is a substantial jump in the use of external consultants. (See Figure 4) Many of these decisions involve the selection of texts, classroom management systems, and the like. It was common in such cases for representatives of book publishers to visit the school or for descriptive materials to be provided by publishing companies for scrutiny by the staff before they made a decision.

This is clearly an evaluative process, though the grist for the evaluative mill is not tests, needs assessment, or input from an evaluator. Such decisions are the one clear example in our data in which there are viable alternatives to be considered in a decision and information is sought out relative to these alternatives. The external consultants provide expert advice that is being used as the basis for making an evaluative judgment between alternatives. INS MATL stands alone in this respect.

Collegial advice also reached its highest level in those decisions involving instructional materials. In these cases it represented another form of expert opinion being brought to bear on a choice. Staff members shared the experiences that colleagues at schools had with the materials under consideration.

STU GRPS. The incidence of test use reaches its highest level in the small number of decisions concerning student grouping patterns (STU GRPS). In these three cases declining test scores were used as a basis for changing the manner in which instruction was being conducted.

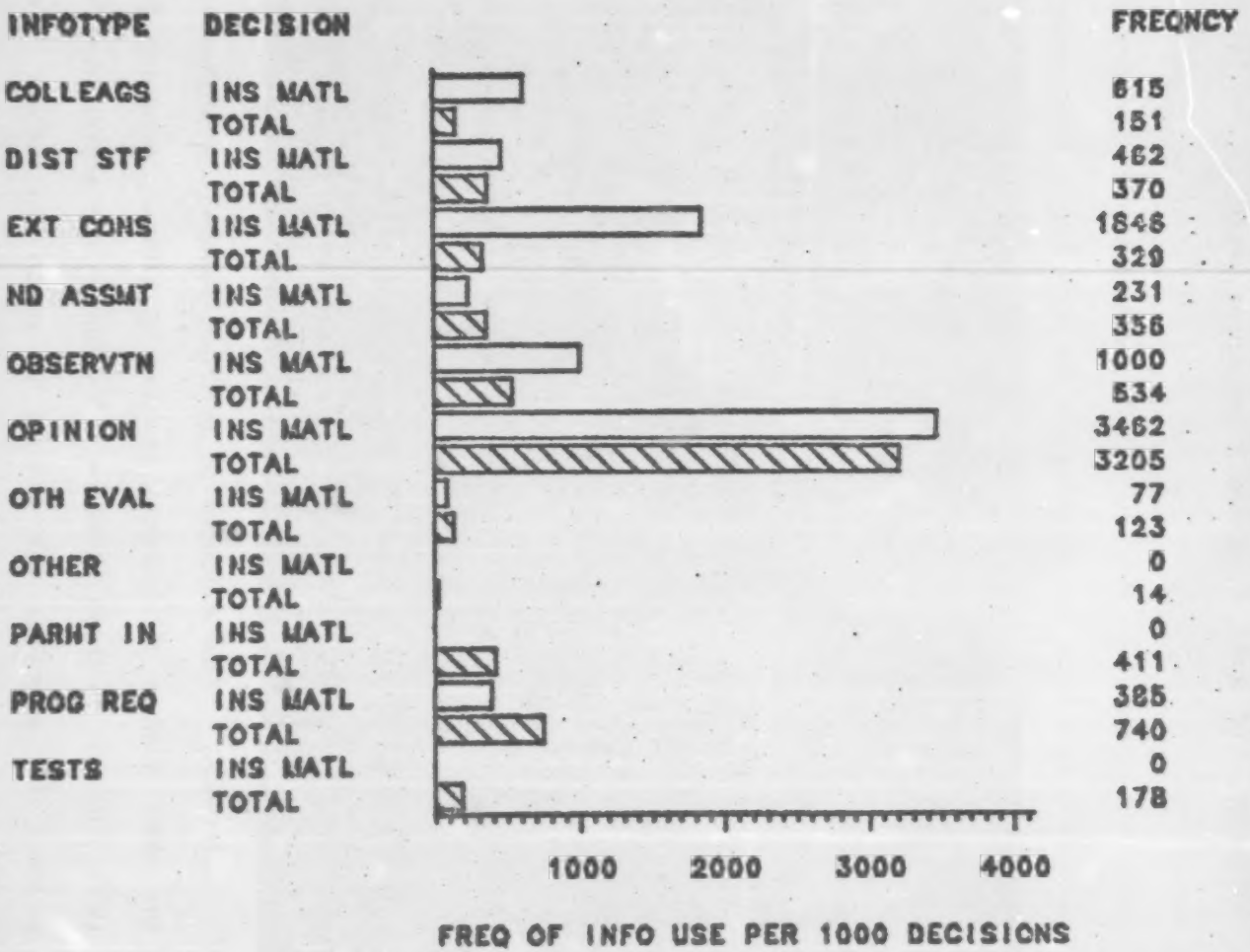

Figure 4: Information Use, INS MATL (N=13) vs. TOTAL (N=73)

BILINGL and NEW PROG. The rest of the decision categories provided no surprises. Program guidelines were featured most prominently in those decisions having to do with bilingual programs (BILINGL) and those relating to the implementation of new programs (NEW PROG). This seems reasonable, as one would expect the greatest reliance on rules and regulations to occur in newer, less familiar program areas. Similarly, district consultants and program personnel from the Metro central office made their greatest input into these same categories of decisions. It seems appropriate that supervisory staff were sought out to help interpret guidelines and develop programs in areas where the school had less experience.

The level of parent input (PARNT IN) was quite high in decisions concerning bilingual programs as well. This observation aligns well with the community-based emphasis of the bilingual programs.

Discussion

This analysis of information use by decision type has confirmed some of the impressions that we developed informally after conducting the interviews.

1. There was a low overall incidence of evaluation use of any type.
2. Needs assessment played a larger role than any of the other types of evaluation activities, but its role was primarily restricted to increasing people's awareness that an action might need to be taken.
3. Similarly, test scores often served as a "flag" warning people that something needed to be done.

4. Only in the case of selecting instructional material was data used to illuminate alternatives. In these instances the data were usually in the form of expert information from outside the school.
5. There was little if any evaluative information used in administrative personnel or staff development decisions.
6. Overall there does appear to be some relationship between the decision type and source information. Looking specifically at evaluation, we found that the required needs assessment activities and review of test data have the potential to promote an awareness of the need for school action in instructional decisions. Evaluative information of the type we examined seems to have little potential use in non-instructional and non-curricular decisions.

PERSONNEL CONFIGURATION VS. INFORMATION TYPE

There were a number of reasons to suspect that some relationship existed between the kinds of information brought to bear on a decision and the personnel who were involved in making it. One reason is derived from organization theory. Hanson (1978) focusses on two types of organizational structure -- bureaucratic and collegial -- and suggests that schools have aspects of both. This is important because these organizational structures have different decision making styles and different patterns of information flow. According to his analysis, the principal's realm is the bureaucratic, while the teacher's realm is collegial. As a result they should show different patterns of decision

making. We were interested in seeing if there were differences in information use when these different personnel groups were involved.

Many other observations from our own analyses also heightened our curiosity about the manner in which personnel affected decision making and, hence, information use. To investigate this relationship we focused on each personnel type separately. For each group we identified the decisions in which they had a high level of involvement and those in which they had little involvement. Then we compared the information profiles between these high and low incidence groups to see if differences existed.

In the discussions which follow we will present only those cases in which substantial differences were found. (Figures illustrating all the comparisons can be found in Appendix G.) The presentation is organized by personnel group.

Teacher Groups (TEACH) We compared the decision sequences in which there was high involvement of individual teachers or small teacher groups with those in which no teacher groups appeared. (See Figure 5) There were no significant differences between the incidence of evaluation use in these two sets of decisions, but there were some differences among other information sources. The most striking difference was in the area of advice from colleagues at other schools. When small teacher groups were involved there was much greater input from colleagues at other schools than when such teacher participation was lacking. Similarly, district consultants were also a stronger force when teacher groups were involved than when the decision process did not involve small teacher groups.

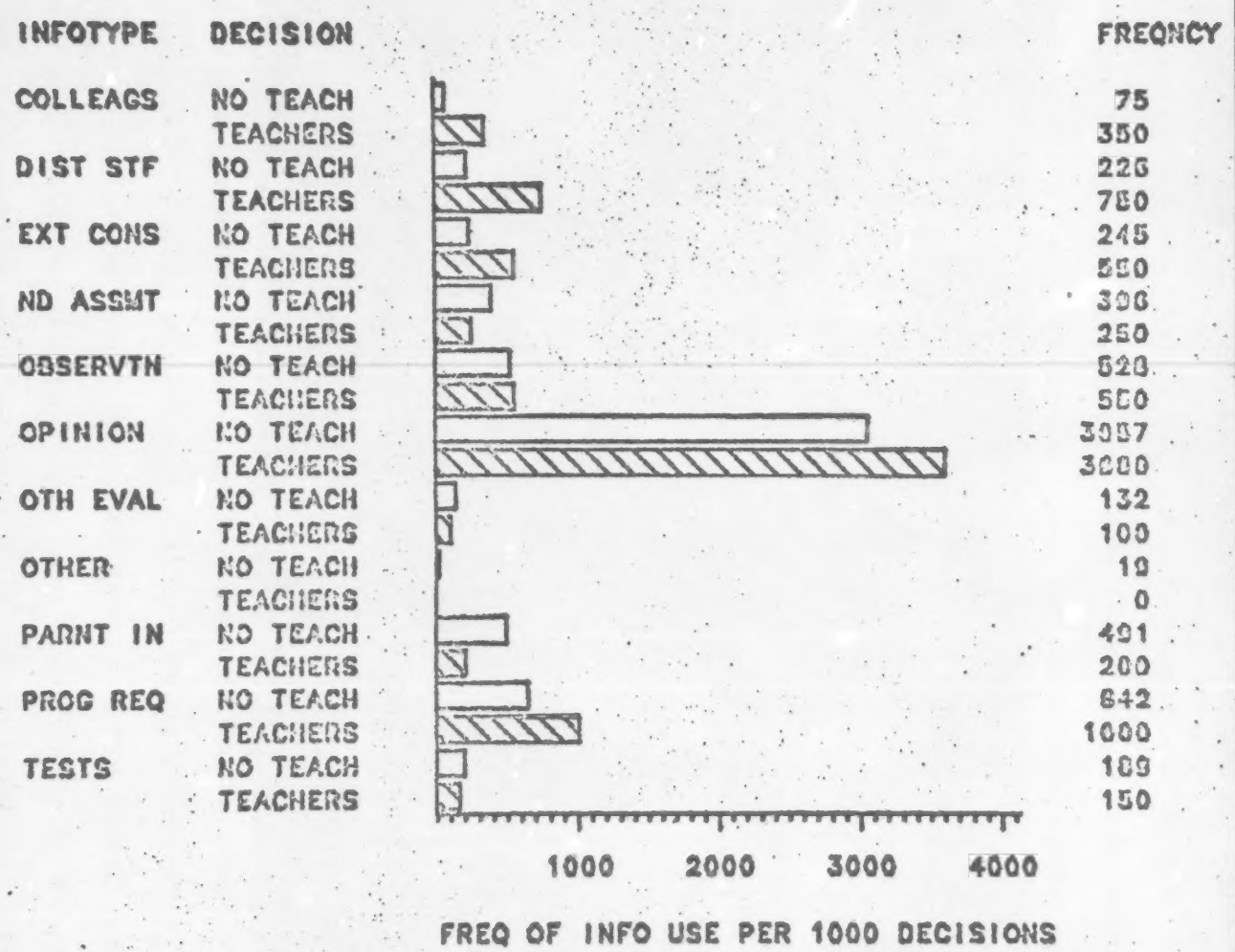

Figure 5: Information Use, TEACHERS (N=53) vs. NO TEACH (N=20)

Administrators (ADMIN) Figure 6 shows the information profiles for administrator involvement subdivided into three levels: none, low and high. There is a consistent growth in the amount of evaluation information that enters the decision process as the level of administrative involvement increases. Needs assessment and tests are cited more frequently in the decisions with greater administrative involvement, and the use of outside evaluation sources goes up somewhat, as well.

Teacher-Administrator Groups (T+AD GPS) We compared information profiles between decisions in which teacher-administrator groups were involved and those in which they were not. (See Figure 7) There is a notable correspondence between the involvement of such groups and the use of evaluative information. Both needs assessment and tests are cited more often when these groups are present. There is also a significant increase in input from external consultants and from district staff when teacher-administrator groups are involved.

None of the other personnel configurations yielded noteworthy differences.

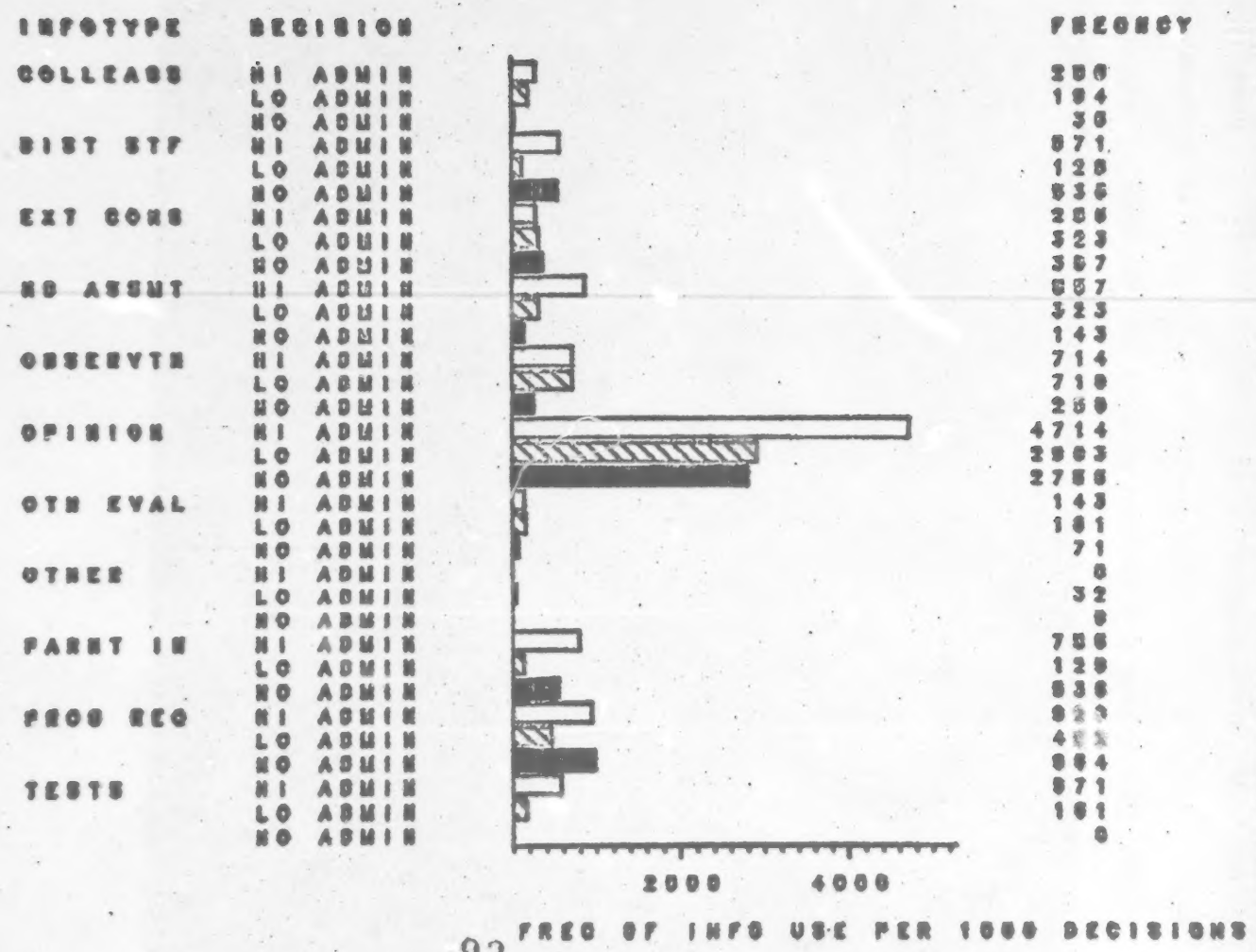

Figure 6: Info Use, NO ADMIN(28) vs. LO ADMIN(31) vs. HI ADMIN(14)

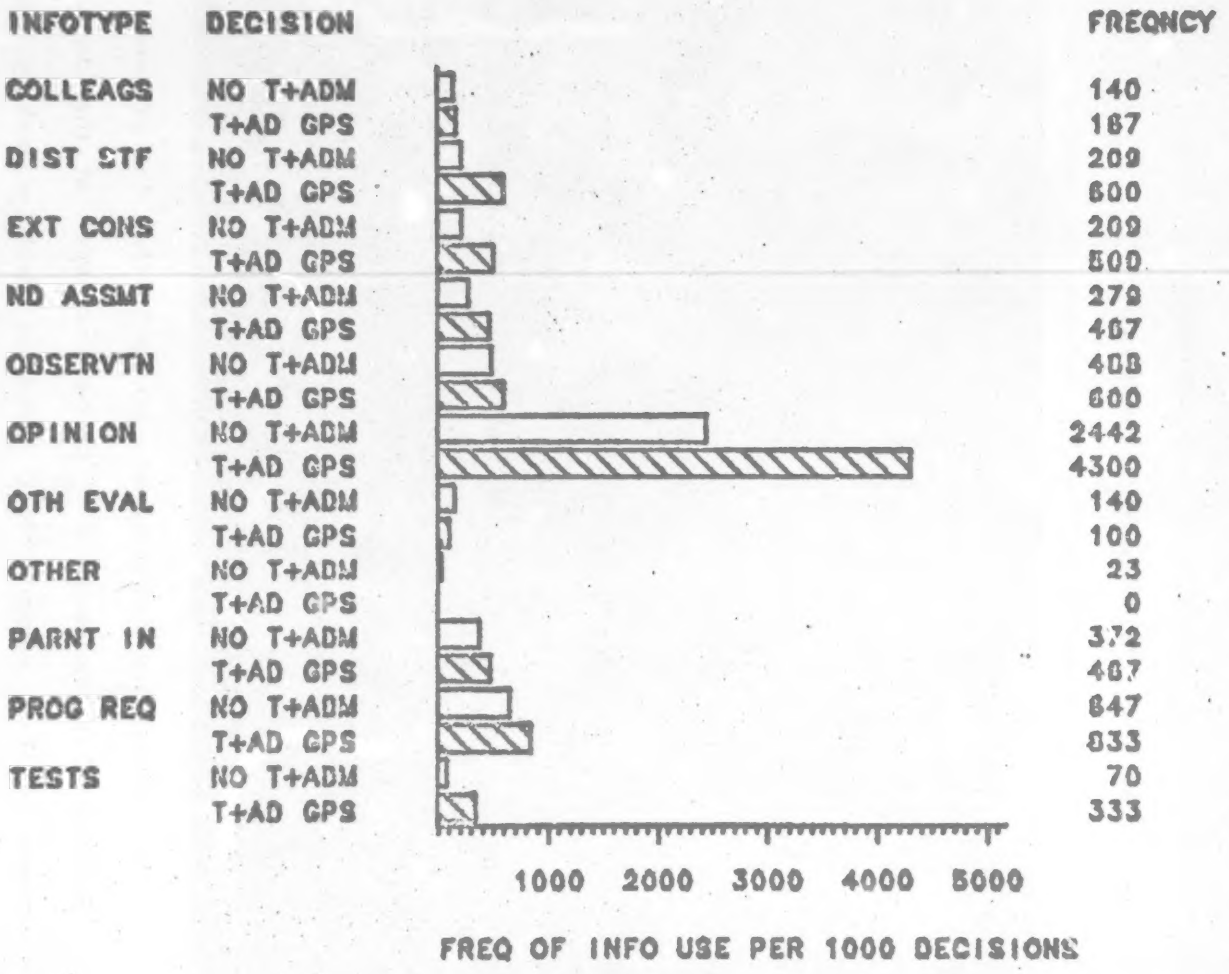

Figure 7: Information Use, NO T+ADM (N=43) vs. T+AD GPS (N=30)

Discussion

One must be somewhat cautious in interpreting these results; the temptation to attribute causality to mere correlation is strong. Mindful of this caveat we offer the following observations:

1. There is no noticeable relationship between the level of participation of teachers acting singly or in small groups and the presence of evaluative information. Similarly, there was little if any relationship between participation of the full staff and references to evaluation.
2. On the other hand, the presence of administrators, whether acting with other administrators or acting in conjunction with teachers in mixed groups, showed a high positive relationship with the level of evaluative data entering the decision process.

Hanson (1978) suggests that decisions in the administrative realm are bureaucratic and involve the exchange of summarized information up and down the chain of command. Evaluative data is this kind of information. On the other hand, decisions in the teachers' realm are more collegial, and this is characterized by greater reliance on personal experience. Our results correspond with this model. However, alternative explanations exist.

One alternative interpretation would be that the presence of administrators increases the "deliberativeness" of the decision process. The administrators formalize decision making, and consequently the process exhibits more careful consideration and rationalized choice.

A third explanation is that teachers have greater constraints on their available time, and do not have the luxury of lengthy deliberations. Administrators on the other hand have more flexible schedules and more time for review and scrutiny of data. As one staff person expressed it:

> I'm sure you must be aware of the fact that a teacher's day is really horrendous in terms of the demands on that teacher's time. (Teachers need free time to think)... Industry has learned this -- I guess we have learned it too, but the price tag makes it prohibitive. I think if we could run one pupil-free day a month, or if we could have two pupil free afternoons a month, or if we had the opportunity to meet together and to interact and to dialogue and share ideas and concerns we would see improvement. But the time constraints are such that it's literally impossible. (13P)

We also note that the teacher-administrator groups consist of "leadership committees" and other specially constituted representative bodies that have a highly rationalized basis for existence. Such bodies, by their very nature, would be more judicial. It is possible that the involvement of such representative bodies insures that a decision will be made in a more rationalized manner.

3. There is little relationship between the presence of other types of personnel -- consultants, parents -- and the level of use of evaluation.

THE RELATIONSHIPS AMONG INFORMATION TYPES

For the sake of completeness we also investigated the relationships among the various information types. We analyzed the data in a manner similar to the procedure used in the previous section -- focussing one at a time on each information type and distinguishing between those decisions in which that type of information played a prominant role and those decisions in which it had only a minor role. These two groups of decisions were compared to see if there were differences in the use of the remaining types of information. Only three of these comparisons yielded any substantial differences. Those were the comparisons based on the variables OPINIONS, TESTS, and COLLEAGS.

Discussion

Examining the comparison based on opinion, we noted that as the amount of personal opinion cited in the decision increased, references to needs assessment and to test results increased as well. One reasonable explanation for this phenomenon is that both needs assessment and test data require interpretation. After examining such data, individuals usually express their opinion about the meaning of the information in light of the issue under discussion.

Testing yielded a more complex pattern. There was a strong positive relationship between references to tests and references to needs assessment -- not too surprising since most needs assessments use test data extensively. There was also a positive relationship between tests and both direct observation and program requirements. On the other hand, there was a negative relationship between tests and both collegial

advice and external consultants. This suggests a differentiation between decisions that were primarily pupil focused and decisions that were primarily program focused. The decisions in which there was higher reference to tests were all drawn from three categories: student grouping, general curriculum and new programs. This seems to correspond more with an inward assessment of local needs than an outward search for advice from others.

Finally, collegial advice was positively related to the use of information from external consultants and negatively related to evaluative information of all types. This seems to corroborate the distinction between "internal" decisions, for which evaluative data play a larger role, and decisions for which external recommendations are sought.

INFORMATION TYPE VS. OTHER VARIABLES

When we reviewed the data tapes, we examined a number of other variables that seemed important based on our initial data synthesis. Each appeared to be related to decision making in some manner, and we wanted to determine if they had a significant impact on the level of information use. The variables were: the number of decision options that were considered (OPTIONS), the length of the decision sequence (LENGTH), whether or not one individual or group had been responsible for creating a strategy around the decision making process (STRATEGY), whether the prompt to action had come from within the school or from outside (PROMPT), and whether one particular group within the school had been strongly identified with the initial idea that a change was needed (GROUP).

We will consider each of the five variables in turn.

Number of options. Overall the number of options ranged from one to four. We were not able to determine how many options had been considered in about a third of the cases. Figure 8 shows the information profiles for those cases in which there was a single option versus those in which there was more than one option. There is a much greater incidence of the use of needs assessment data and the use of external consultants in the multiple option circumstances. Again, we must be extremely careful in interpreting these results that we do not derive causal inferences from mere associations. While a plausible argument could be made that the presence of multiple options leads to greater reliance on these two types of data, the causal link might actually be the other way around. For example, external consultants might be the ones who suggest new options. Yet these are not the only two reasonable interpretations; a third variable might be causing the variation we observed. This would be the case if, for example, the difficulty of the problem was causing the staff to seek outside help and generate more new options of their own. Finally, of course, there may be no causal linkage between the two variables at all.

The broader knowledge of the decision context derived from the interviews provides us with more information to bring to this question, though we may still be unable to establish any stronger interpretation.

Length. The length of the decision process was determined by counting the number of distinct steps that were related by the respondent. Figure 9 shows the different information use patterns between those occurrences in which there were only one or two steps before the final decision was reached and those with a longer deliberative process.

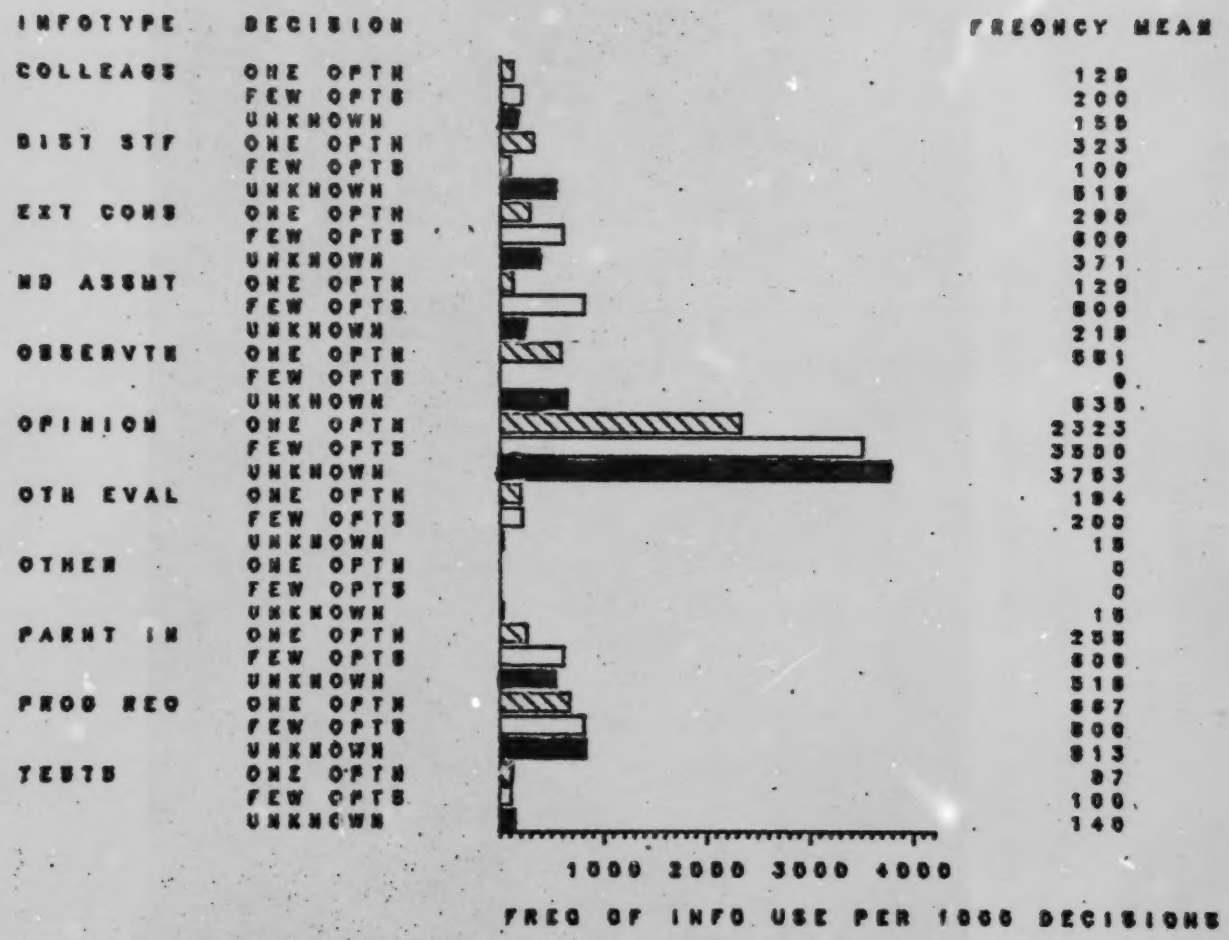

Figure 8: Info Use, ONE OPTN(31) vs. FEW OPTS(10) vs. UNKNOWN(32)

As the length of the decision process increased, the incidence of evaluative information grew. Similarly, the incidence of almost every type of information increased as the process lengthened, with the greatest increase occurring in input from parents.

Strategy. Sometimes people create a strategy or set of steps for coming to a particular decision. For each decision we determined whether or not such a plan had been established and, if so, who was responsible for establishing that strategy. In Figure 10 the information profiles comparing the levels of the the STRATEGY variable are displayed. We compared those instances in which a staff person or group of people collectively took responsibility, those in which the principal was responsible, and those in which no one established a strategy for action.[12] The differences between the first two categories were not very great. While the incidence of needs assessment cited among those decisions in which the strategies were established by a group is larger than those decisions in which the strategy was coordinated by the principal, the situation was just reversed for other types of evaluation. The total of all three evaluative sources of information is about the same for the two groups. However, we do find a difference when we compare these totals with the decisions in which there was no strategy. The level of use of many of the information types is less in the latter case.

---

[12] This was often the case. The decision either evolved organically or followed an existing standard operating procedure that was part of the regular school routine.

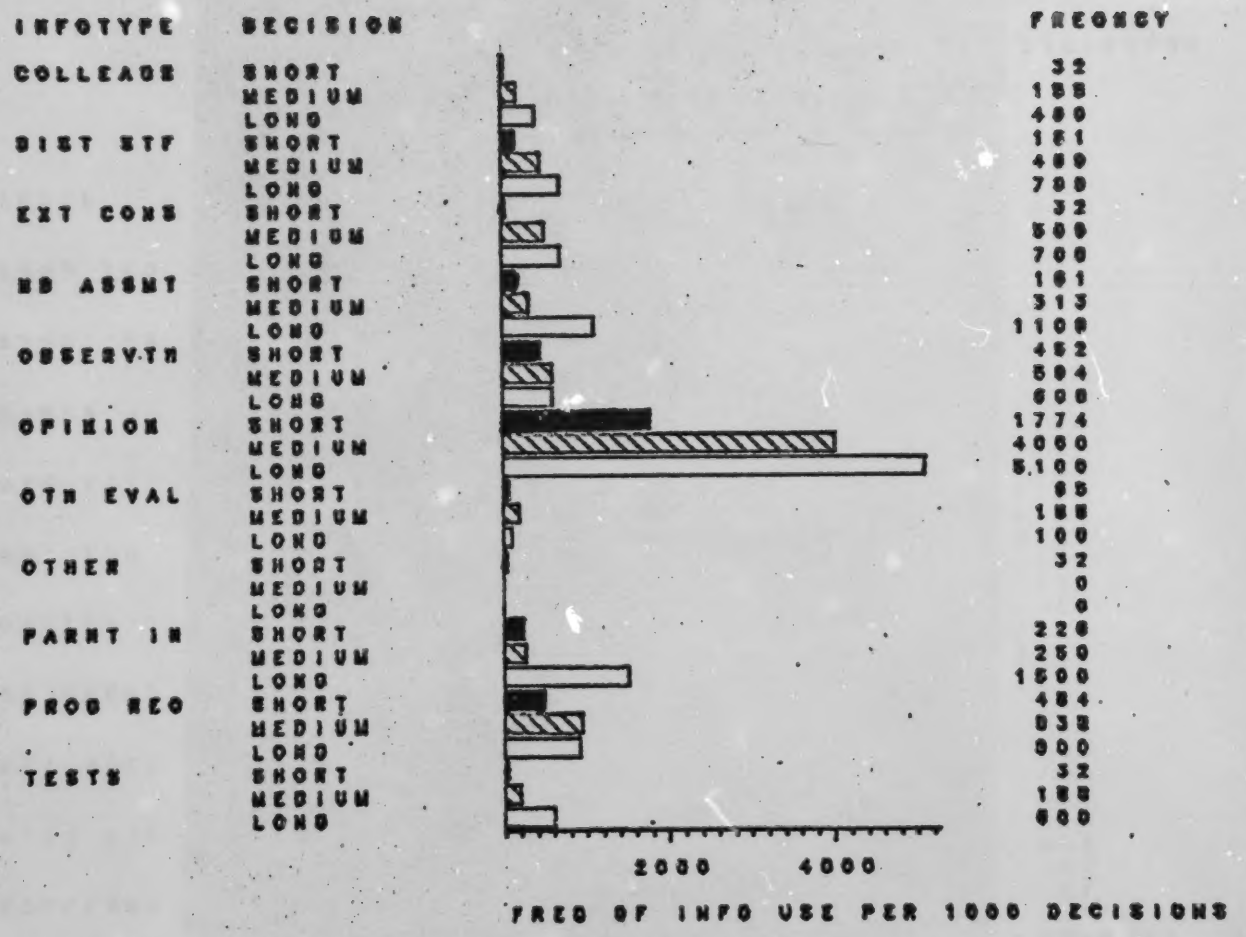

Figure 9: Information Use, SHORT(31) vs. MEDIUM(32) vs. LONG(10)

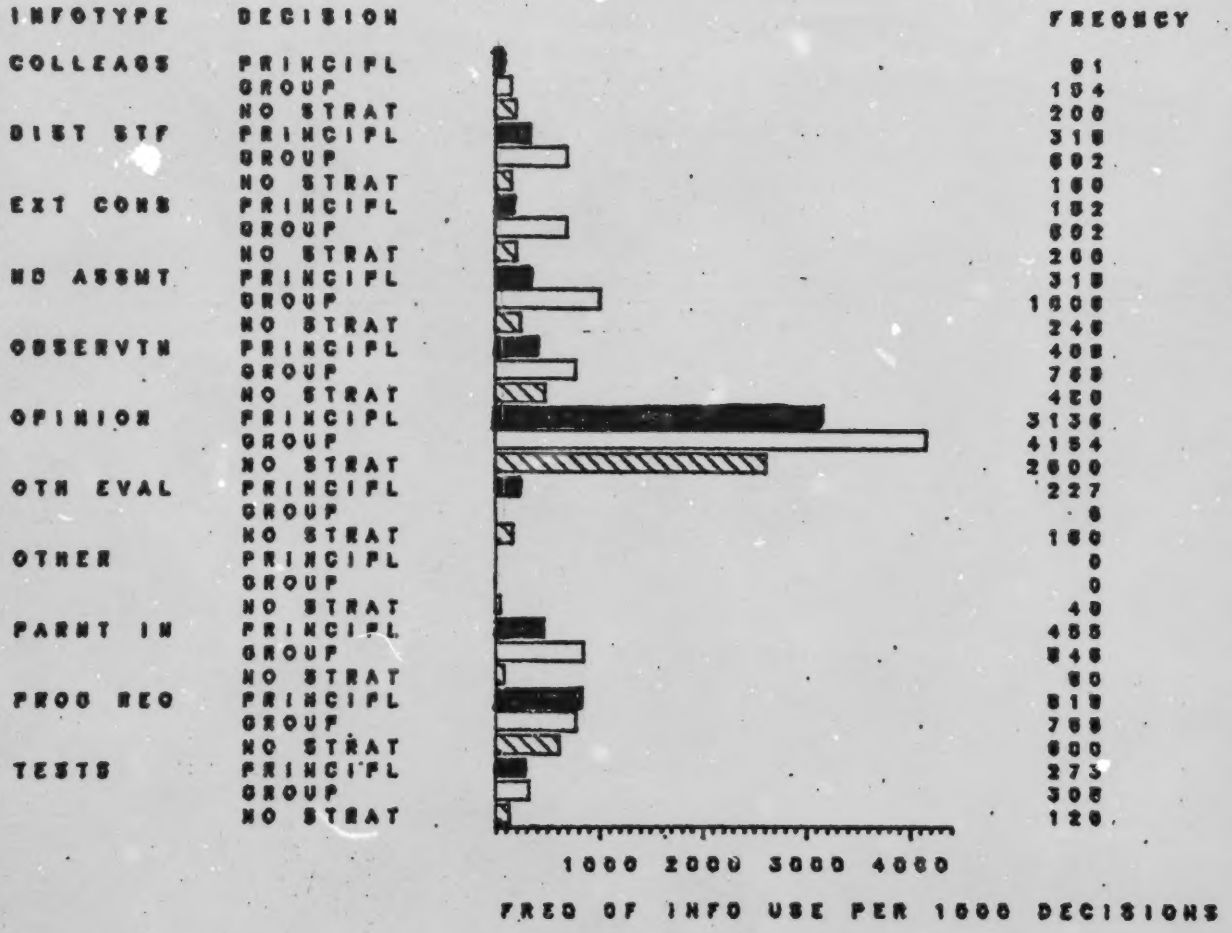

Figure 10: Information Use, PRINCIPL(22) vs. GROUP(13) vs. NO STRAT(25)

Prompt. Every significant occurrence began with a recognition phase in which an initial prompt which was noticed by a person or persons within the school. We were able to distinguish between those prompts that arose within the school itself and those that emanated from outside. Figure 11 compares the information profiles of internally and externally prompted occurrences. There are no significant differences between the level of evaluative information used in these two sets of decisions. Understandably, there is much greater reference to program guidelines and district consultants in cases in which the prompt was external, while there is much greater mention of direct observation when the prompt to the decision came from within the school.

Groups. When the prompt was *internal* we looked to see whether a particular individual or group of individuals was strongly identified with a particular change. In Figure 12 we compare the information profiles among those decisions in which different in-school groups were strongly identified with a particular change. (Recall, there were many decisions in which no such group was apparent, so the sample we are reviewing is smaller.) There is a marked difference in evaluative information, especially needs assessment, between issues identified strongly with teachers or administrators *alone* and those identified strongly with a mixed group. The same pattern also holds for parent input. There do not appear to be differences between the groups in any of the other types of information.

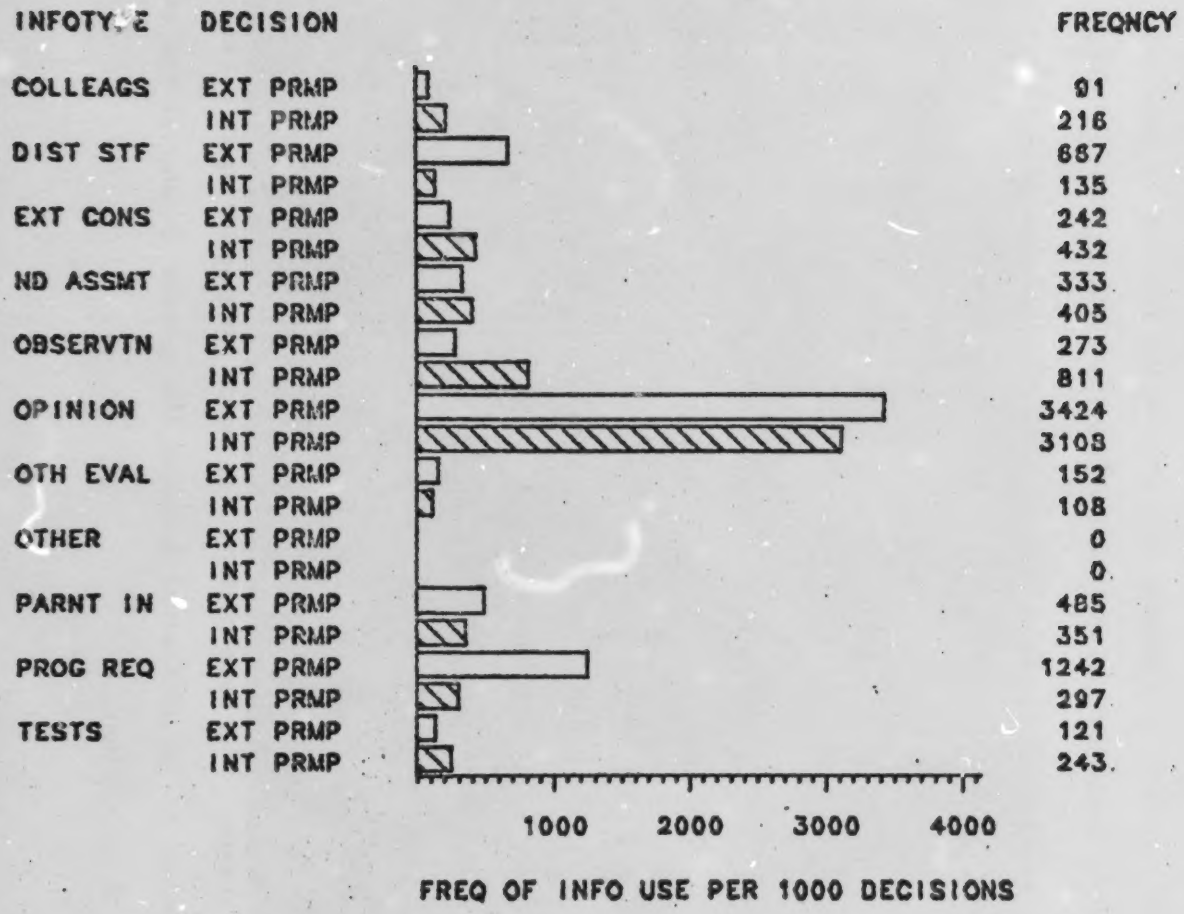

Figure 11: Information Use, INT PRMP(N=37) vs. EXT PRMP(N=33)

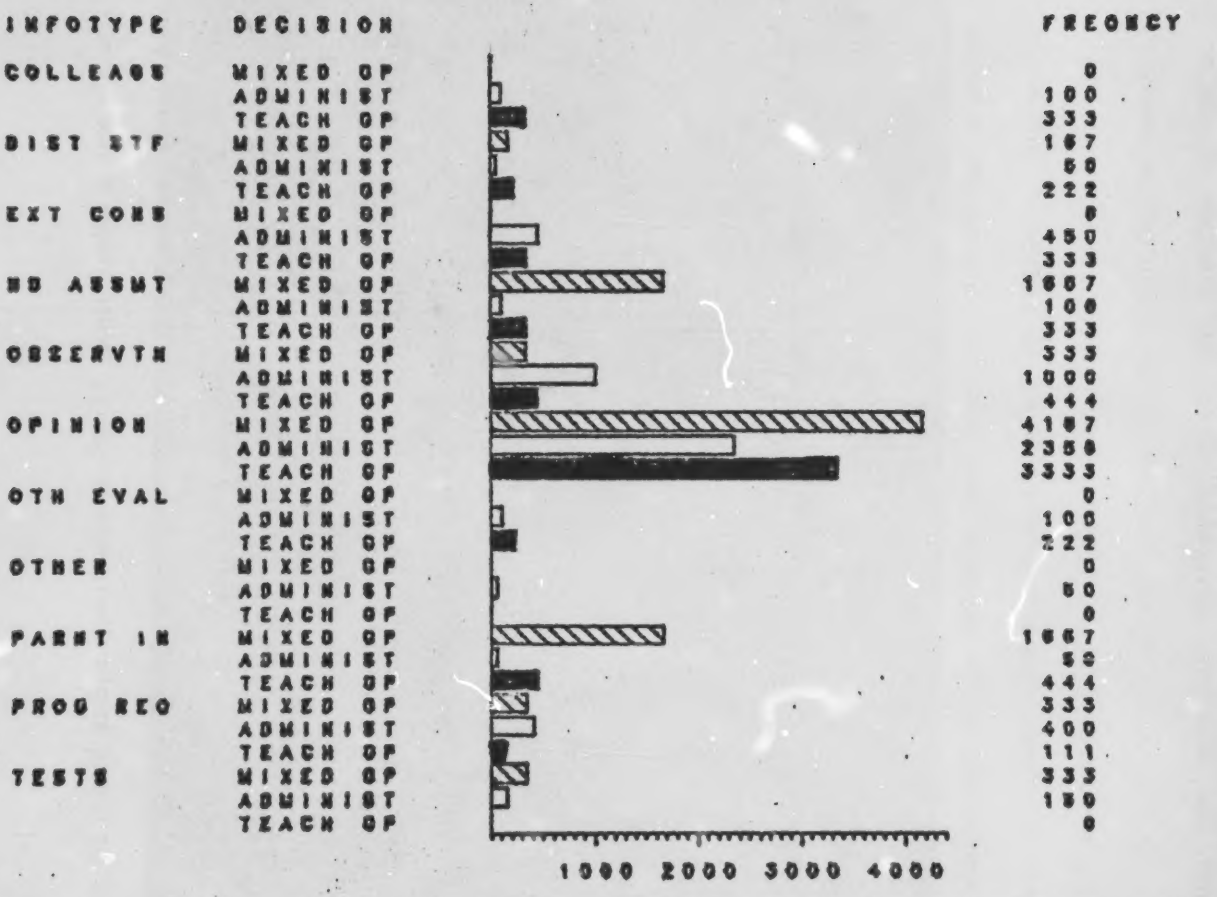

Figure 12: Info Use, ADMINIST(20) vs. TEACH GP(9) vs. MIXED GP(6)

Discussion

To understand the apparent difference when the number of options increase we note that many of the multiple option occurrences involved a choice among textbooks.[13]

1. In most of these cases representatives of textbook publishers were contacted which explains the greater reliance on external consultants. It was usually the case that teachers deliberated and expressed their choices among the options, thus the increased incidence of personal opinion makes sense as well.
2. Length seems to have a great effect, but this is not really a variable subject to external manipulation. It is not surprising that longer decisions involved more information (the decision may have been prolonged by certain parties insisting that more information be considered), nor that the greatest increase was in the level of input from parents. Parent input is channeled through School Site Councils and School Advisory Committees, and these bodies were only involved in the more elaborate and formalized decisions. Daily decision making is of little concern, as only large-scale, school-wide program development issues are brought to the parent councils for comment. Such actions, e.g., the annual program application, are lengthy, multi-stage procedures.

---

[13] Keeping in mind the caveat that association by itself does not imply causality, we can still interprete these results in light of our total knowledge of the phenomena under study.

3. Strategy is also important. Many educational researchers have commented on the importance of key individuals in school decision making. We suggested in an earlier working paper that the principal was such a person, and his or her leadership style was a prime determinant of evaluation use. What Table 10 seems to suggest is that it is not so much the principal who determines evaluation use *per se*, but any individual or group of individuals who step in to take the lead in coordinating a decision. The main differences were not between the principal-led occurrences and the group-orchestrated ones, but between these two categories and those occurrences in which there were no groups that created a strategy or plan of action. The combined evaluative total (TESTS + ND ASSMT + OTH EVAL) for the first two categories is about the same, but this is markedly greater than the evaluative total for the latter set of occurrences.

4. Prompt seems to have little impact on evaluation use, though the distinction between internally-prompted and externally-prompted decisions makes a difference in other types of information use.

5. The differences due to the GROUP variable are somewhat more difficult to understand. One way to interpret the strong increase in reference to needs assessment when mixed groups of teachers and administrators are strongly identified with an issue, is to remember that needs assessment often acts as a "cause" itself, not just as secondary data. That is, the data provided to the whole school as a result of the needs assessment process may point out an area that requires attention. The parent

committees are usually involved in the needs assessment process as well; consequently, we are not surprised to find greater parent input based on the same information that motivated the administrators and teachers to opt for change.

Chapter 5

DECISION PATTERNS ACROSS ALL FOUR PHASES

INTRODUCTION

One of our initial goals was to characterize school level decision making processes in a manner that would allow us to look for recognizable patterns. We characterized the school's actions related to each significant occurrence in four phases. To this point we have focused our attention on the decision making phase because that was where the greatest potential for evaluation utilization lay. In the first half of this chapter we look more closely at the other three phases. After offering some overall comparisons between the phases we will discuss each phase in turn. Then, in the second half of the chapter we will describe the prototype decision sequences we developed to summarize decision making patterns.

ANALYSIS OF THE DECISION PHASES

Comparison between the Phases

In Tables 4 and 5 we have summarized the pattern of information use and the breakdown of personal configurations that were reported in each phase. (The data are reported as the number of citations per 1000 decisions, as they are in all the figures in this report.) We will discuss each phase in turn.

TABLE 4

Frequency of Information Use in Each Phase

Frequency per 1000 Decisions

Information Type	Recognition	Decision Making	Ratification	Dissemination
OPINION	667	3205	1000	23
PROG REQ	402	740	147	0
OBSERVTO	235	534	88	23
PARNT IN	98	411	118	0
DIST STF	98	370	0	0
ND ASSMT	157	356	29	0
EXT CONS	29	329	0	0
TESTS	137	178	118	0
COLLGAGS	0	151	0	0
OTH EVAL	39	123	0	0
OTHER	69	14	0	23

Tables 4 and 5 illustrate quite clearly that the bulk of the interactions took place in the decision making phase. The number of personnel involved and the level of information use were both many times greater in this phase than in any other. On the other hand, the relative frequencies among the types of information and personnel display their own patterns. Ignoring the dissemination phase (for there was essentially no information involved in the dissemination of the decision) the relative balance of the information types is similar from one phase to the next. However, there is much more variation in the rela-

TABLE 5

Personnel Configuration in Each Phase

Frequency Per 1000 Decisions

Personnel Configuration	Recognition	Decision Making	Ratification	Dissemination
ADMIN	598	863	441	0
STAFF	118	849	265	535
T+AD GPS	88	493	59	0
TEACH	88	342	59	70
PAR+STFF	59	219	176	70
PAR+AIDE	10	68	59	395
CONSULTS	20	41	0	23

tive magnitude of various personnel types between the phases. This will be clearer as we discuss each individual phase, but some overall comments seem warranted at this juncture.

The balance of information types in the decision making phase has been reviewed extensively in previous chapters, and the predominance of opinion noted in the decision making phase holds in the recognition and ratification phases as well. However, its relative role vis-a-vis the other information types is somewhat lessened in the recognition phase. This makes sense because there was less of a role for opinion in recognizing factual changes and events (such as new program guidelines, changing school populations, and low test scores) than in deciding how to respond to these prompts. The three evaluative information types occur with differing relative frequencies in the three different phases

we are discussing as well. The relative role of tests, needs assessments and other evaluation is greater in the recognition phase than in the decision making phase or the ratification phase.

The balance between different personnel types shifts more dramatically than the balance between information types as we compare phases. Administrators dominate the recognition phase,[14] while there is more balance in the decision making phase between the administrators and the other members of the professional staff. Ratification is primarily the function of the administrator or the full staff, with some involvement of parent advisory groups, while dissemination of a decision goes mostly to the full staff, to the parents and aide group or to the parent council. There will be more to say about the relative balance of personnel and information in the following sections when we analyze each phase individually.

Recognition

In the recognition phase we captured the earliest reported identification of a need for school action. It was not always easy for our respondents to make this judgment, because many of the significant occurrences that were described materialized gradually over time. For example, many schools in our sample experienced growth in the percentage of their student population who were from Hispanic background, but this was a slow, incremental process. It was difficult to identify the point at which someone recognized the need to make changes to accommodate these students. In fact, in most cases an external

[14] It appears that a stimuli is not officially recognized as important until it is legitimized by an administrator.

reporting or planning cycle prompted the formal recognition that the gradual change had reached a threshold that required action. In this case events such as the annual program application process, the filing of the district racial/ethnic survey or a school-wide needs assessment crystallized the staff's view of their situation.

Reviewing Table 4 one notices that personal opinion was the predominant type of information cited in the recognition phase and that direct observation also was cited quite frequently. This tends to support our belief that 'recognition' was sometimes a personal and subjective phenomenon which depended on a key individual's view of a changing scene.

Program guidelines are one criteria that is used to determine if school action is required. The large number of references to program requirements suggests this was a common mode of action. Typical of such situations was the case in which the number of LES/NES students reached certain levels and instructional changes were required by law. The high incidence of PROG REQ references also reflects the fact that, many times, changing requirements themselves became the prompt for action. The creation of the supplemental instructional program for racially isolated schools was such a situation.

The level of citations for needs assessment and test data reflect situations in which evaluative data drew attention to a potential problem or area of improvement.

There are no surprises in the distribution of personnel in the recognition phase. Table 5 shows that the administrators, who are responsible for coordinating the school's overall program, were most often the people who recognized the need for change (or who received the notifi-

cation that official changes were being made in guidelines or procedures). In addition, there were a fair number of significant occurrences which were first vocalized in full staff meetings, and there were *some* instances in which each of the different personnel configurations was responsible for recognizing a need for action.

Ratification

The ratification phase was described as an official review stage in which some person or persons were given an opportunity to comment on a decision tentatively agreed upon by another group. Our group discussions uncovered three basic ratification sequences: the full staff confirmed a decision made by a committee or group, the principal 'signed off' on a decision made by teachers or the full staff, and the parent/teacher committee ratified a decision made by the professional staff.

The data in Table 4 confirm this picture; the personnel involved in ratification are administrators, the full staff and parent-staff groups, in that order. In contrast to the decision making phase in which parent input was primarily indirect, we do find direct parent and aide participation in the ratification stage. Here parent deliberative bodies such as school-site councils and school advisory committees were frequently involved in "signing off" on plans development i. the school.

The information types cited in the ratification phase also add credence to our earlier conceptualization. The only type of information that is referenced to a significant degree is personal opinion. We are not witnessing a complete recapitulation of the decision process with all

arguments and points of view, but rather an abbreviated review of the final choice in which a group is given an opportunity to express their own ideas.

(It should be noted that we were told of no instances in which a decision was 'vetoed' in the ratification phase.)

Dissemination

We did not anticipate that many types of information were required in the dissemination phase and we found exactly that. The process that was described was one in which decisions made by groups or by administrators were disseminated to the full staff, or decisions made by school professionals were disseminated to the parents. That is essentially all that is depicted in Tables 4 and 5

Discussion

This model provides a more complete picture of the full decision process that occurs in the schools, though these three phases hold limited interest for our study. The one important element is the evidence that evaluation -- in the form of test scores and needs assessment -- is directly involved in the recognition of many problems. In fact, it contributed roughly 15% of the total number of information citations in this phase.

PROTOTYPE SEQUENCES (BY DECISION TYPE)

The final phase in our analysis was to examine the complete decision making process to see if similar patterns existed among decisions of the same type. If generalized decision prototypes could be found they would be powerful tools for investigating evaluation use in school actions and might also suggest ways to enhance evaluation use.

We were somewhat successful in this effort, extracting prototypic decision sequences for certain cases but not for others. The search itself was illuminating. We diagrammed the decisions sequences that had been described to us and found that diversity predominated over similarity. On first inspection it seemed that every sequence differed in some small manner from every other. In fact, even when we aggregated our units of analysis to the personnel and information groups used in the previous discussions, the differences often outweighed the similarities. Though somewhat disappointing, this diversity is in itself one of the important findings of our research.

Equally important amid this widely varying set of decision patterns were some similarities. We were able to identify some generalizable prototypes. In the sections that follow we will describe these decision prototypes and give specific examples from our study.

First, a brief word about notation. We will use the previously defined categories to label steps in the decision prototypes. People, rather than information, dominated the descriptions of decision making that were provided by our respondents. Consequently, we used the personnel group involved in each step as our defining element and the predominant information sources as a secondary element of the notation.

Also, for completeness, we consider the recognition phase, the decision making phase and the ratification phase although most prototypes contain only two of the three phases. Ratification did not occur too often, and the form usually varied from case to case. (A slash is used to indicate the end of the recognition phase and the beginning of the decision making phase; two slashes separate the decision making phase from the ratification phase.)

For example, a prototype might be designated in the following manner:

ADMIN(ND ASSMT) / ADMIN, TEACH, STAFF(ND ASSMT, OPINION)

This example indicates a decision having both recognition and decision phases. The following sequence of actions might have occurred in a situation that was described by this prototype. Initially, the principal or other staff person recognized the deficiencies in the school's reading program when he/she conducted the annual needs assessment. The decision phase included several steps. First, administrators discussed the reading program among themselves and came up with some of their own ideas. Then they shared the scores collected during the needs assessment process with the grade level chairman (teachers) and brought them into the discussion of the school's response. All agreed that the problem was the school's departmentalized reading program. Many felt that it was not working well and that teachers wanted to return to self-contained classrooms. The next step in deciding what to do was to discuss the issue at a full staff meeting the following week. Here, all the teachers agreed that something had to be done to improve the scores -- all agreed that the best thing was to switch back to self

contained classrooms. This decision was made. Throughout the decision process the predominant information had been the data collected during needs assessment and the opinions and observations of the staff themselves.

Our prototype does not attempt to capture every single bit of information used in the decision, nor to display every contributing interaction between school personnel. Rather, it is a global model of the important steps in the decision process and the most salient pieces of information that were brought to bear on the problem at hand.

This sample prototype indicates one of the chief roles of evaluation we observed in the schools -- identifying the need for change. It is not the only role that evaluation plays, nor do all decisions evolve in this manner. However, a certain class of decision resemble this problem, and it is a useful tool for characterizing those situations.

The various prototypes we were able to identify seemed to fit better with particular types of decisions. As a result, we have organized the presentation of decision sequence prototypes according to the types of decision used initially to classify significant ocurrences.

General Curriculum Prototype The prototypic decision involving general curriculum guidelines was as follows:[15]

ADMIN(PROG REQ, TEST) /

ADMIN, STAFF, T↑AD GPS(TEST, OPINION)

---

[15] We will describe this prototype in great detail and provide a lengthy example in order to familiarize the reader with the notational system. In subsequent prototypes the description and example will be more succinct.

The key elements of this prototype are the need for action recognized primarily by an administrator with critical elements in this recognition being the program regulations and the recent set of test scores or observations. Generally, the administrators discussed it among themselves, enlarged the discussion to include some sort of representative teacher group or sampling of staff, and finally brought in the entire staff who was ultimately responsible for deciding the course of action. Key elements in this process were the opinions of the personnel, their likes and dislikes regarding the suggestion put forward by the administrator, and the program guidelines themselves. Evaluation was not brought to bear on the consideration of alternative courses of action, but served merely to signal at the beginning of the sequence that something needed to be done.

Such a decision occurred at school number 3:

> Example: Reading scores had been low for the last few years and the assistant principal wanted to do something about it. Because he believed that the teachers' instructional behaviors were not as well organized and planned as they could be, he developed a management plan for the reading program that he wanted to implement in all classrooms. It corresponded more closely to the goals that they had set out in their program application and to the overall district curriculum guidelines. He discussed his ideas with the principal who gave him his approval to broach the subject with the faculty. The principal did not want to order the change, and hesitated to force it upon the staff. However, the assistant principal showed a lot of enthusiasm and got permission to present the idea at the executive committee meeting. The executive committee was a representative teacher group that would meet with the administrators on a regular basis. They were somewhat cool to the plans suggested by the assistant principal. They thought they were unworkable, extra wieldy and awkward, and they suggested a number of changes. During the next two weeks the assistant principal made some changes in his original management plan outline, got the reluctant approval of the executive committee and presented it at a full staff meeting. The staff were not completely convinced, either, that this was the right approach. They didn't like being told how they should go about managing their classrooms.

However, they recognized that the scores had been declining and that they would have to take some actions to make some attempt to improve the situation. They suggested one or two other changes and reluctantly agreed to implement the new plan.

Most of the decisions in the GEN CURR category differed from this prototype in some manner. However, the prototype captures a common thread suggested by the whole group of decisions considered together.

Instructional Material Prototype. Most significant occurrences involving changes in instructional materials accrued over a long period and had no prototypic recognition phase. However the decision phases had some marked similarities. The instructional materials prototype is as follows:

T+AD GPS, TEACH, T+AD GRP, (OPINION, EXTCONS, COLLEAGS)

// STAFF (OPINION)

The typical instructional materials decision involved the selection of new textbooks. Usually some dissatisfaction with existing texts had been brewing for a long, but indefinite, period. An executive committee or a representative teacher committee usually was seeking out information from text publishers, from the district, and from colleagues at other schools. Alternative texts were displayed at the school and publishers' representatives often were invited to make presentations. The teacher committee actually decided which text to purchase but the entire staff was called in to approve the final decision after hearing a report. The staff usually relied on their own opinions about the books in making their choice.

What we note in the instructional materials prototype is a very minimal administrative participation. Administrators usually played a role in establishing a framework for the process, delegating resource teachers or certain classroom teachers to gather information and review available materials. Ultimately, however, the full staff chose the instructional materials and administrators usually accepted their recommendations.

New Program Prototype. The new program prototype reflects a situation that may be unique to the district we studied. Additional funding was given to certain schools to provide more after school teacher services. The model of the prototype is as follows:

ADMIN (PROG REQ) / T+AD GRP, STAFF (OPINION, PROG REQ)

A description given by a staff person at school number 16 exemplifies this pattern:

> Example: The principal was notified by the district office that the school was eligible to receive special funds, and was also provided with the requirements that must be met in order to receive the funding. He shared this information at the next leadership committee meeting, and also informed this committee that he was going to let the teachers decide how they would organize their after school hours. Each teacher would have to prepare a brief written statement indicating what kind of activities would be going on in class to supplement the regular instruction, but the choice would be left up to the individual teachers. This information was shared with the whole staff and the teachers made their choices based on their own personal preferences and the limitations that were set by the requirements of the program. The selections were all reasonable and the principal didn't feel it necessary to veto any of them.

Not all of the occurrences in the New Program category related to the use of additional funds for the special program schools.

Parent Involvement Prototype. Parent involvement decisions seemed to be made exclusively by administrators. Although there were few decisions in this category, the following prototype seemed to apply:

ADMIN(OPINION) / ADMIN, PAR•STF(OPINION)

The principal usually did not initiate the action under discussion, but rather an assistant or a coordinator who had been delegated the responsibility for parent involvement did. When the idea had been refined and a course of action identified, parents were consulted to insure that it would meet with wider approval. Typically the members of the school site council were consulted or some parents who were active in the schools and who were frequently on the school grounds were brought into the discussions. When this group of parents and administrators agreed that the alternative was a good one, the final decision was made.

Personnel Prototype. There were very few personnel decisions described among our significant occurrences. Those that were described adhered quite closely to this prototype:

ADMIN (OPINION) / ADMIN (OPINION)

The personnel category was limited to decisions that would normally be considered as falling within the purview of the school principal -- which is what we found when we analyzed those cases.

Bilingual Prototype. The bilingual program decisions varied widely. However, the key elements are captured in the following prototype:

T•AD GPS(PROG REQ, DIST STF, ND ASSMT) /
ADMIN, TEACH, PAR•STFF(OPINION,PROG REQ)

Most of the decisions in the bilingual category arose out of the changes in the school population. The disparity between the bilingual program requirement (particularly the district's Lau plan) and the situation at the individual school created a need for action. A member of the district staff who was familiar with the problem involved in the bilingual program usually was involved in monitoring the school's efforts and pointing up deficiencies. District consultants had the most expertise in how to meet the needs of bilingual students with the limited resources available. Their input about viable and acceptable options primarily determined the course of action to take. This decision sequence reflects more district input than any of the other prototypes we looked at so far.

Another new element in this prototype is parental participation. The district did not have enough certified bilingual teachers, and often met the needs of students through the use of bilingual classroom aides who were drawn from the local parent community. Moreover, program requirements dictated advisory participation of the school advisory committee.

The prototypic bilingual decision sequence started with the school administrator's recognition of the problem. The problem usually involved having to make some instructional adjustments to serve a larger number of certified teachers. The bilingual coordinator for the bilingual teacher group was often involved in planning how the changes would be made; these plans were communicated to the parent representatives who had a chance to comment and the whole staff was also given opportunity to participate in the decision.

Other Significant Occurrences. We were not able to extract a representative prototype for the other six categories of significant occurrence. In each instance, we were unable to find a pattern for the occurrences that fell into the six categories: 1) STU GRPS, 2) STF PERS, 3) STFF DEV, 4) SML INST, 5) EVAL REL, 6) MISC. These categories were represented by too few cases or had too much diversity for us to identify a prototypic model for the actions taken.

Discussion

Our attempt to develop prototypes for the decision types described to us is both illuminating and frustrating. In particular, we note:

We were able to characterize a typical decision sequence in half the categories of significant occurrences. These prototypes indicate the course through which the decision took place, and the critical information sources that were brought to bear. In this regard, they provide a very efficient shorthand for discussing a complex phenomena.

On the other hand, some of the decisions defied our attempts to characterize them in this manner. They shared a common subject, but they proceeded in very different ways and used different kinds of information. One thing that this might suggest is that the commonality among the subject matter was not as great as we thought. Our classification scheme could have inaccuracies which only showed up when we tried to diagram the decision. Some of the categories were more uni-dimensional than others and thus, perhaps, more amenable to the development of a decision prototype. A closer look at the six categories for which we were not able to develop decision prototypes lends credence to this interpretation.

The decision prototypes capture the order in which personnel were involved, but they do not capture the *influence* that one group had on another. This is an important drawback to using the prototypes as a model for improving evaluation utilization. They shed some light on which kinds of information were important, but not what influence they had. They capture *what* happens but not *why*.

To a large extent the same groups of personnel were involved in most decisions. If we focus on the curriculum areas (ignoring staffing decisions and personnel matters), the personnel who appear in prototypes look very similar. Most of the decisions involve administrators, groups of teachers and administrators and the whole staff at some stage in the process. The main differences were not in *who* was involved in the decision, but how they influenced one another and what information they brought to bear. The decisions that involved major issues, ones that school staff deemed significant, usually involved all the different groups of personnel of the school. The difference seems to lie more in the kinds of information brought into the debate than in which personnel.

The typical decision sequence is short, lasting just two or three steps. This may be because the school environment is very hectic; there is much to do and little time to do it. As a result we saw very few instances of elaborate, deliberative processes and lengthy considerations. The standard procedure seemed to be to make the best possible decision with the information at hand or readily available. There were no instances in which the process looked like a theoretical decision making process in which alternatives were generated and information (including evaluation) brought to bear on those alternatives.

Chapter 6

CONCLUSIONS

In this chapter we will summarize the more important results of this study, consider some possible refinements and discuss the implications of this research for evaluation practice and for future research on evaluation utilization.

SUMMARY

The User Interview Survey achieved the goals that had been set for gathering, categorizing, and analyzing information about evaluation use among elementary school decision makers. Although the findings of the survey are recorded in detail in Chapters 3, 4, and 5, we will bring some of the results together in summary so that broader, more general patterns can emerge more clearly.

We asked our respondents to identify "significant occurrences", and their selection is noteworthy in itself. By far the most commonly described occurrences involved general issues of curriculum and instruction. (These did not include, however, any direct intervention in instructional practice within the classroom.) Thus, while the respondents had broad discretion to interpret the notion of a "significant occurrence" in any manner they chose, they generally agreed that instructional and curricular issues were the most important.

One of the first observatons we made after reviewing the list of significant occurrences was that the majority of the events were not initiated by the school personnel themselves. Over half of the significant occurrences had their genesis outside the school, as reactions to federal, state, and district actions or to community changes. Schools spent most of their time *reacting* to events rather than initiating them. The overall picture of schools that emerged was one of institutions with a desire to undertake constructive efforts to improve instruction in the face of multiple external demands on time and resources. Not surprisingly, there was also some resentment about these constant pressures from outside. This anger and frustration must be kept in mind when thinking about ways to improve evaluation utilization.

Looking at the questions of the use of evaluation, we found that school decision makers did not frequently rely upon evaluation when they made decisions. Instead, they acted most heavily on the basis of personal belief and opinion. Program guidelines and regulations were given the second greatest amount of attention.

At this point we must digress from the discussion of specific findings to talk about the nature of the results we obtained. Predominately, the generalizations we were able to draw were valid only for certain types of decisions, for particular phases in the decison process or for certain types of evaluation. In fact, one of the most important findings of this study was that overall generalizations about school decision making or evaluation were not possible; definable patterns of behavior or interaction only were found to be applicable for particular circumstances.

The universe of generalization for most of our conclusions is not school decision making but school decision making *of a particular sort*. For example, personnel decisions operated differently than decisions involving the establishment of general curricular guidelines. Similarly, the notion of evaluation in the aggregate is too broad for useful generalization. There were different uses for needs assessment than for the assistance afforded by the Metro evaluation consultant. Finally, we found that it was useful to subdivide the decison process into a number of smaller phases, and that different relationships held in these different phases. The use of evaluation, in particular, differed between the recognition phase, the decision making phase and the ratification phase. Thus, to summarize, we were able to make important distinctions between different conditions and to produce a number of conditional generalizations.

The nature of these conditional generalizations becomes more evident when we further consider some of the findings relating to information sources. Though evaluation data played a very small role in the decision making phase, they played a much larger role in the recognition phase. Both needs assessment and testing were useful in identifying areas that needed school attention. Consequently, we must qualify our initial pessimistic assessment of evaluation utilization. While it is true that evaluation was not greatly present in the full decision making cycle, it played an important role in one part of that process -- problem recognition.

Needs assessment, in particular, was a type of evaluation which was mentioned frequently in the recognition phase by our respondents. We

found that needs assessment helped school staff identify areas that required attention. By evaluating the status of the school program on an annual basis in a form that was familiar and in a manner that involved the staff directly, needs assessment had a sizeable impact. It helped the school staff recognize some of the successes and failures of their program. These findings confirm the conclusions of Brown and Braskamp (1980) that needs assessment was used to stimulate interest, raise new issues and serve as a basis for future evaluation activities.

Within the decision making phase of the sequence there were differences in the use of evaluation data depending upon the decision type. Evaluative data (primarily tests and needs assessment) were more likely to be used in certain types of decisions. Evaluation did *not* enter into administrative decisions, staff development decisions or personnel decisions. However, it was important in curricular decisions and in decisions involving the bilingual program. This observation makes good sense. Evaluation of the kind that was described to us is not germane to purely administrative actions, nor is it particularly relevant to most staff development and personnel decisions. What little evaluation use we found in the decision making phase was concentrated in curricular and instructional decisions, and this is somewhat heartening.

The study also examined whether there was a differential impact when different personnel were involved in decision making. First, we noted that most of the decisions that were described to us involved the *entire* professional staff at one point or another. A multistage process usually occurred in which different individuals or clusters of people were involved at different points in time. We do not mean to imply that

decision making was democratic. In fact, the role of one group or individual was usually dominant while the role of another was more limited. However, most of the significant occurrences that were described to us were accessible to the whole professional staff at some point. In contrast, only rarely did the district consultants become directly involved in school level decision making, and personnel from the Evaluation and Testing office were never mentioned. Paraprofessionals and parents also had only limited direct involvement in decision making, though their input was often conveyed indirectly through teachers or administrators. Decision making in these significant occurrences could almost be viewed as a family affair among the school's professional staff with little direct participation by "outsiders".

Secondly, the level of evaluation use was related to which personnel groups participated in the decision process. In particular, the presence of administrators acting singly or in groups with teachers was related to increased evaluation use. It may be that administrators had more time to devote to considerations of evaluation, that they were more familiar with the information that was available, or that they had better training and a stronger commitment to data based decision making. Whatever the case, the level of evaluation utilization increased in those decisions in which the administrators participated, either alone or in groups with teachers.

In the next stage of our analysis we tried to develop decision prototypes for each type of significant occurrence. This was potentially the most important part of our analysis. Not only did it reinforce one of the major conclusions about evaluation utilization to arise from the study

-- the distinction between the use of evaluation in the recognition phase and in the decision making phase -- but it may have the greatest long term implications for increasing evaluation utilization.

We drew two important, yet seemingly contradictory, conclusions from the attempt to identify decision prototypes. First, identifiable similarities exist in decision making. We were able to characterize distinct prototypes that captured the essential common elements of the decision process for six classes of decisions. This is a major accomplishment. Such prototypes can be a valuable tool for understanding evaluation use, and, as we will see, for developing prescriptions to increase such use under different decision conditions. Second, identifiable prototypes do not always exist. In six other types of decisions the differences outweighed the similarities, and we were unable to develop prototypes of common action patterns.

What does this mean? For one thing, the phenomena under study were enormously complex, and any attempt to aggregate by focussing on similarities must ignore a multitude of individual differences. Furthermore, as suggested earlier, the categorization system itself might have been responsible for some of the heterogeneity in certain groups of decisions. Beyond this, there still may be wide differences, and we cannot determine from this study how great they are. These findings suggest that further study to verify the similarities we captured in the six prototypes seems warranted, and more detailed study of the other kinds of events is certainly called for.

This concern for independent validation of the prototypes is a good introduction to a discussion of potential refinements to this study that

might be undertaken. Following that discussion we will review some of the implications for evaluation practice that might be derived from our findings.

REFINEMENTS

As noted above, further verification of the applicability of decision prototypes in other situations seems like a valuable exercise. While these prototypes are valid for Metro district, local conditions (particularly the administrative structure) vary from district to district, and this may in turn affect decision procedures. While we believe conditions at most elementary schools are similar in essential ways, this issue does warrant further investigation.

Hearkening back to a commment we made in Chapter 4, this study tells us a lot about *what* occurred in the decisions we studied, but much less about *why* it occurred the way it did. The "what" is valuable in itself -- we learned a lot about the decision making process -- but it also leaves a great deal to investigate. Why did opinion predominate? There are any number of possible explanations for this fact -- opinions are usually salient, familiar, trustworthy, immediate and credible. How is opinion formed? We did not thoroughly investigate the important elements that went into the formation of these opinions -- evaluation could well have been one of the factors that subtly shaped people's attitudes.

The following example illustrates another issue that might be addressed as a refinement to this study. A colleague, who worked in Metro district for many years, tells us that a form of evaluation exists

which is highly relevant to personnel decisions. Each teacher applicant is given an entry examination in his or her primary subject area, and these scores become part of the person's personnel file. These data are usually scrutinized by the principal before any hiring takes place.

We were surprised that these scores were never mentioned in the personnel decisions that were described to us. On closer review we noted that none of these decisions involved simply hiring a new teacher. They involved increasing the amount of time provided by a specialist already employed, or shifting staff among different jobs. Under these circumstances it was unlikely that the principal would refer back to personnel records.

Our colleague offered another explanation. In her view, such things as the entry test are so commonplace that they might not be mentioned. They become part of the "background noise" that is filtered out because it is so familiar. Our respondents might simply have failed to mention the test scores because they were common knowledge within the district and therefore not prominent in their recollection of the event.

The concern raised by this discussion is not the use of the evaluative instrument in personnel decisions *per se*, rather there is another, more important implication. This example points out the inherent limitation of retrospective acounts of an event as complex as decision making. Although our respondents indicated that they were able to recall the details of the events to their *own* satisfaction (and we checked that they had little to add when we sought field verification), we have no way of knowing how much of the "background noise" was filtered out in both cases.

The critical point is that there are limits to the amount of information that can be obtained in an hour long retrospective interview. We tried to insure the accuracy of these reports through extensive cross-validation procedures, and we are convinced that no contradictory statements or blatant falsehoods were included in our data. Nevertheless, our results are constrained by the accuracy of our respondents' memories and the sensitivity of their perceptions.

The reports contained considerable detail -- enough to make the data analysis itself a challenge. However, the more familiar we became with the events under discussion, the more we recognized the value that could be derived from even more detailed reconstructions. Refined observations and data collection procedures directed toward underlying causes could yield a fuller recounting of this aspect of the events that transpired, and thus shed more light on the reasons underlying the patterns we observed.

IMPLICATIONS

This study has a number of important implications both for future research on evaluation utilization and for evaluation practice in the schools.

Research Implications

Up to now evaluation research has failed to distinguish between types of decisions, and the assessment of evaluation's impact is inaccurate when this distinction is overlooked. By identifying distinct decision types, this study begins to balance the assessment, and such distinctions should be included in any subsequent research.

Similarly, most evaluation research has focused on the decision-making stage and has ignored the other stages in which evaluation may play an important role. By concentrating on the manner in which evaluation was acted upon, past research has ignored evaluation's more subtle influence at other stages. This study's recognition of stages in the decision making process further corrects previous over-generalizations about evaluation's role.

When we carefully examined the stages in decision making and diffentiated between decision types, we were able to derive conditional generalizations about evaluation use and make more precise statements about particular types of decisions. For example, because it focused on the end product of the decision, most past research failed to perceive evaluations's importance in identifying the very problems which were being addressed. This study suggests, therefore, that further research should include examination of the decision making process in its broadest sense, including recognition, decision making and all other phases.

We initiated the study to look at interrelationships in a large cross-section of decision making, knowing full well that our method of analysis would limit our ability to infer causes. Now that this analysis is complete we would like to see efforts to achieve greater understanding of the "whys". It seems worthwhile to expand this inquiry to include a larger number of decisions, to allow for lengthier interviews or even first hand observations of decision making, and to include data from the rest of the school staff. Without doubt teachers should be included in any subsequent investigations of this sort.

In fact, certain relationships may never be uncovered without observing decision making *as it is taking place*. For example, any study that would hope to ascertain why opinion predominates in decison making (a question we raised earlier) would probably want to rely on direct observation of decision processes. Similarly, one could only hope to understand the role of evaluation in opinion formation (another of our earlier concerns) through a review of opinion formation over time.

This suggests that an ethnographic study of school decision making would be valuable. The investigator should remain at the school for an extended period of time and observe first hand significant occurrences of the type described to us. Such contextually sensitive research would be an important supplement to the broad cross-sectional investigation undertaken in this study. It could begin to fill in some of the missing "whys" that were only alluded to in our findings.[16]

On considering the implications of this study for future research we also note the success of the data reduction and analysis techniques we employed, and recommend that future research in this area consider similar approaches. These were not simple tasks. Data validation procedures employed on this project were uncommmon to most qualitative research. The multiple analysis and aggregation procedures were also quite novel. Finally, the use of the computer to make comparisons among the quantified variables from qualitative data yielded a variety of important insights -- easily justifying the expenditure of time and energy.

---

[16] An earlier ethnographic study which was part of CSE's Evaluation Use Project did provide many of these kinds of insights (Daillak, 1980). However, that study focused on the district office evaluator and not the school.

Finally, we laid the foundation for a classification scheme that may refine the discussion of evalution utilization and prove essential in further research. The two-dimensional breakdown of school decisions by type and phase had clear utility in this study -- the prototypes derive directly from this structural model -- and it should be explored further. The decision type by decision phase matrix that emerged from this study appears to be a useful organizational tool for studying evaluation utilization at the school level.

Practical Implications

The study has a number of implications for evaluation practice. We will highlight some of these and discuss what might be done to improve evaluation utilization in light of these results.

First, we should emphasize the importance of the context in which evaluative activies are conducted. Much of what we observed was a function of the structure and operating procedures of the district as well as the school. For example, the role of the evaluation consultants was stipulated very clearly by the district. Daillak (1980) noted the degree to which their activities were circumscribed, and noted even that some people went outside their official duties to interact in alternative, informal ways. This is all to say that one cannot consider evaluation utilization in the school and ignore the impact of the district administrative structure. The external pressures we noted above give ample evidence of this fact. In fact, Daillak, Alkin & Stecher (1980) noted that administration itself seemed to be a much more salient concern than achievement at the local level, and this observation is

confirmed in the present research. Under these circumstances prescriptions for improving school site evaluation utilization must involve both district as well as local factors.

The clearest instance of evaluation use -- the use of needs assessment in the recognition phase -- points up the importance of local involvement and familiarity in the evaluation utilization process. More attention is given to data that are locally generated. Such data are more familiar and they have greater credibility at the school site than information that is communicated from outside the school. In addition the personnel at the site have a personal investment in needs assessment information because they are actively involved in its collection. Recalling the almost family-like exclusivity of most important decision making supports the notion that, to be useful, information must have a local basis. In contrast, there was an almost complete lack of input from the Metro evaluation consultant, and there was only minimal attention given to evaluation in other forms. School personnel proffered negative reactions to external mandates and directives that emanated from the administrative hierarchy. In fact, often the evaluation consultants acted as the enforcers of such requirements. An implication to be drawn for improving evaluation utilization is that the responsibility for initiating and gathering evaluative data related to significant occurrences must be shifted to the local site.

This notion was recognized by administrators in Metro district. Daillak (1980) noted that the E & T office initiated an effort to establish on-going planning and monitoring committees within each school. Unfortunately, the implementation of these local evaluation committees

was not given high priority, and individual evaluation consultants were given great flexibility in terms of the amount of emphasis they placed on the effort. As a result, ongoing planning and review never became a reality. We saw little evidence of this effort in our interviews. Only once or twice was such a committee even mentioned. We think this shift of responsibility for evaluation would go a long way toward improving utilization.

Yet, our research suggests that certain functions might be successfully carried out by such a local committee, while others could not. The distinction we would make is between information that is collected to serve external reporting functions, and information that can fill a local need. David (1978) noted that most Title I evaluation was carried out for reporting purposes only, and these are precisely the kinds of externally mandated activities toward which the respondents in our study reacted most negatively. In order for local evaluation efforts to contribute to school improvement they must be motivated out of local concerns and must serve local needs. Needs assessment (though mandatory) has shown that it can inform local decision making in a useful manner, and it is accepted to the degree that it does so. Other evaluative activities will have to pass this same test in order to achieve increased utilization.

This distinction between information for external mandates and local site needs argues for a separation between compliance and reporting activities on the one hand, and evaluation for local decision making on the other. While the EUP has earlier argued that evaluators should adopt a consultative role, and we still feel that this approach has the

greatest potential for increasing evaluation utilization, we would suggest that it would be impossible for a district evaluation consultant to spend enough time at an individual school site to handle all the data that might be useful for local decision making. The capability for performing evaluation must be shifted to the local schools themselves. The district evaluator could be an instrument of this change, undertaking training and technical assistance functions in a consultative manner, but probably could not handle the on-site responsibility for such evaluation.

Our analysis of decision making has other practical implications. We can use the breakdown of significant occurrences and the decision prototypes to make predictions about the kinds of issues that are likely to arise during the course of the year, and the kinds of information that are likely to be useful in addressing those issues! For example, we know the types of significant occurrences that happened most frequently in Metro district and the typical manner in which many of them were addressed. With this information school staff could generate evaluative information that would be useful in a particular decision. The end result of such a procedure is that evaluation could play a much greater role in many important, and seemingly predictable, school decisions.

When we began this data analysis, school decision making appeared to be haphazzard. It seemed to be dominated by unpredictable changes and events rather than by careful planning or reasoned review of information. The school level decision makers, who provided this viewpoint, addressed significant issues as they occurred without preliminary plan-

ning. They found little use for existing evaluation and relied instead on personal opinion to make decisions.

The results of this study suggest that there is some identifiable order under this chaotic facade, and that the existing pattern of decision making can be altered. If information on significant issues, such as that derived from this study, can be fed back into the system to illuminate that order and provide guidelines to help local schools develop relevant evaluation, then the role played by evaluation in local decisions can be increased. In our view, the kinds of analyses conducted as part of this research have the potential to increase evaluation utilization at the local level.

BIBLIOGRAPHY

Alkin, M.C., Daillak, R., and White, P. *Using evaluations: Does evaluation make a difference?* Beverly Hills, CA: Sage Publications, 1979.

Alkin, M.C., Kosecoff, J., Fitz-Gibbon, C., and Seligman, R. *Evaluation and decision making: The Title VII experience.* CSE Monograph Series in Evaluation (No. 4). Los Angeles: Center for the Study of Evaluation, University of California, Los Angeles, 1974.

Alkin, M.C., Stecher, B.M., Daillak, R.H. Interview Survey of Users: Interim Report. *NIE Deliverable, Evaluation Use Project,* Los Angeles: Center for the Study of Evaluation, University of California, Los Angeles, 1980.

Aronson, S.H., and Sherwood, C.C. Researcher versus practitioner: Problems in social action research. *Social Work,* 1967, *12,* 89-96.
Brown, R.D. & Braskamp, L.A. Summary: Common themes and a checklist. *New Directions for Program Evaluation.* San Francisco: Jossey-Bass Inc., Number 5, 1980, 91-98.

Daillak, R.H. Evaluators at work: A field study. *NIE Deliverable, Evaluation Use Project,* Los Angeles: Center for the Study of Evaluation, University of California, Los Angeles, 1980.

Daillak, R.H., Alkin, M.C. & Stecher, B.M. "Evaluation in a period of transition." In L. Datta (Ed.), *Local, state and federal change: Evaluation Issues.* Beverly Hills, California: Sage Publications, forthcoming.

David, J.L. *Local uses of title I evaluations (Research report EPRC 21).* Menlo Park, California: Educational Policy Research Center, SRI International, 1978.

Griffiths, D.E. *Administrative theory.* New York: Appleton-Century-Crofts, 1959.

Guba, E.G. The failure of educational evaluation. *Educational Technology,* 1969, *9,* (5), 29-38.

Hanson, E.M. *"organizational control in educational systems a case study of governance in schools."* Paper presented at the meeting of the American Educational Research Association, Toronto, Canada, March 1978.

Lortie, Dan C. *School Teacher.* Chicago: University of Chicago Press, 1975.

Mann, J. "The outcome of evaluative research." In C.H. Weiss (Ed.), *Evaluating action research: Readings in social action and Education*. Boston: Allyn and Bacon, 1972.

Meyer, J.W., & Rowan, B. Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 1977, *83*, (2), 340-363.

Miles, M.B. *Common properties of schools in contest: The backdrop for knowledge utilization and "school improvement"*. New York: Center for Policy Research, 1980.

Patton, M.Q., Grimes, P.S., Guthrie, K.M., Brennan, N.J., French, B.D., and Blyth, D.A. *In search of impact: An analysis of the utilization of federal health and evaluation research*. Minneapolis: University of Minnesota, 1975.

Patton, M.Q. *Utilization-focused evaluation*. Beverly Hills, CA: Sage Publications, 1978.

Patton, M.Q. *Qualiltative evaluation methods*. Beverly Hills, CA: Sage Publications, 1980.

Rodman, H., and Kolodny, R.L. Organizational strains in the reacher-practitioner relationship. *Human Organization*, 1964, *23*, 171-182.

Weick, K.E. Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 1976, *21*, 1-19.

Weiss, C.H. "Utilization of evaluation: Toward comparative study." In C.H. Weiss (Ed.), *Evaluating action programs: Readings in social action and education*. Boston: Allyn and Bacon, 1972. (Reprint of a paper presented at the meeting of the American Sociological Association, Miami Beach, Florida, September, 1966.)

Weiss, C.H. Knowledge creep and decision accretion. *Knowledge, creation, diffusion, and utilization*, 1980, *1*, (3).

APPENDICES

Appendix A: Framework for Studying Evaluation Utilization

Appendix B: Interview Guide

Appendix C: Interview Topic Description

Appendix D: Interview Summary Form

Appendix E: Coding Form

Appendix F: Frequency of Information Use: Comparisons Between Each Type of Significant Occurrence and the Total Sample

Appendix G: Frequency of Information Use: Comparisons Between Different Personnel Categories

APPENDIX A

Framework for Studying Evaluation Utilization

Cat. 1	*Preexisting Evaluation Bounds*	
	Property 1.1	School community conditions
	Property 1.2	Mandated bounds of an evaluation
	Property 1.3	Fiscal constraints
	Property 1.4	Other nonnegotiable requirements
Cat. 2	*Orientation of the Users*	
	Property 2.1	Questions or concerns about the program
	Property 2.2	Expectations for the evaluation
	Property 2.3	Preferred forms of information
Cat. 3	*Evaluator's Approach*	
	Property 3.1	Use of a formal evaluation model
	Property 3.2	Research and analysis considerations
	Property 3.3	Choice of role
	Property 3.4	User involvement
	Property 3.5	Dealing with mandated evaluation tasks
	Property 3.6	Rapport
	Property 3.7	Facilitate and stimulate the use of information
Cat. 4	*Evaluator Credibility*	
	Property 4.1	Specificity
	Property 4.2	Changeability
Cat. 5	*Organizational Factors*	
	Property 5.1	Interrelationships between site and district
	Property 5.2	Site-level organizational arrangements
	Property 5.3	Other information sources
	Property 5.4	Teacher and staff views
	Property 5.5	Student views
	Property 5.6	Costs and rewards
Cat. 6	*Extraorganizational Factors*	
	Property 6.1	Community influence
	Property 6.2	Influence of other governmental agencies
Cat. 7	*Information Content and Reporting*	
	Property 7.1	Substance
	Property 7.2	Format
	Property 7.3	Information dialogue
Cat. 8	*Administrator Style*	
	Property 8.1	Administrative and organizational skills
	Property 8.2	Initiative

APPENDIX B

Interview Guide

Name of Interviewer:_____ School Code:_____

Respondent Code:_____

INTERVIEW GUIDE Title:_____

0. Introduction
   - Who we are interviewing
   - Why ("uses of information in special programs")
   - Confidentiality
   - Appreciation
1. Description of Specially-Funded Programs (Consolidated Project)

2. Duties & Responsibilities

3. "Significant Occurrences in the Life of the Program"
   - Changes (personnel, goals, materials, attitudes, etc.)
   - Rejected Alternatives
4. Factors Affecting Identified Occurrences
   - Description/History
   - Different Influences
   - Resolution Process
5. Role of Evaluation in Identified Occurrences

6. Role of Evaluation in General
   - Administrative Level
     (Within-school, District sponsored, PQR & mock review)
   - Description
   - Influence on Action & Attitudes
   - Factors Affecting Impact
   - Improvement?

(Repeat if appropriate: 5. Role of Evaluation in Identified Occurrence)

Additional Comments

APPENDIX C

Interview Topic Description

Interview Topic Description

(Training Document)

Introduction to the Study

The purpose of this research is to determine the role that information, particularly evaluation information, plays in school level program decisions. It is difficult to ascertain the relative importance of evaluation information directly. Asking about evaluation tends to bias the respondents' recollections towards just those situations in which they did consider information from evaluations. Instead, the school-level decision makers will be asked to identify significant occurrences in the life of the school programs. The situations they select will be analyzed to determine the factors that affected their beliefs and actions. Among these factors may be evaluation.

Hour long interviews will be conducted with school-level administrators, who might be users of evaluation information. These will not be structured interviews with rigid protocols, but naturally evolving conversations guided toward certain carefully selected topics. The topic guide is outlined below. The precise wording of questions asked by each interviewer will not be predetermined, rather it will evolve within the topic framwork as part of the natural conversational style of the interviewer. Similarly, the exact ordering of questions will be an interactive function of many factors, including, for example, the focused or diffuse quality of the respondent's answers, etc.

Model Introductory Remarks

Hello, my name is _____. We are interviewing elementary school administrators to investigate the ways they use different types of information in school planning and administration. We are particularly interested in schools with specially-funded, supplemental programs.

I can assure you that everything we say in this interview will be strictly confidential, and any reports that are written will be completely anonymous. If you do not object, I would like to tape record our conversation. It allows me to capture your thoughts correctly, and makes our work much more accurate. However, if at any time you would like to stop the recording for a moment, please indicate that to me and I will turn off the machine.

I would like to start by asking you for a brief description of the specially-funded programs here at _____ school.

Description for Interviewers

Topic Area 1: Specially-Funded Programs in the School

A basic knowledge of the nature and scope of the specially funded programs in each school is necessary to understand the context in which decisions occurred. Initially, only a very general description will be sought; specific details will be elaborated as part of the subsequent inquiry into selected events and occurrences.

Model Opening Question: I think the easiest place to begin is with a description of the program here at _____ school. Can you give me a very brief description of the programs you have here as part of the school's Consolidated Project?

Topic Area 2: User's Position and Responsibilities in the School

We also need to know each respondents duties and responsibilities in the school. In particular their administrative relationship to the school's special programs will be important. At the outset a very general description will suffice. Details will be obtained as specific decisions are investigated later in the interview.

Model Opening Question: Can you give me a general description of your job and what your duties are with respect to the programs you just described?

Topic Area 3: Significant Occurrences in the Life of the Program

This is a crucial question, for the respondent's answer will determine the situations on which the bulk of the interview will focus. Ideally, each respondent will be able to recall significant program decisions in which they participated. Realistically, however, the evolution of a school program is more a matter of

incremental change than formal "decision" events. Thus, each school administrator will be asked to identify two or three events that they believe were "significant occurrences in the life of the program(s)". Subsequently, the interview will focus on these occurrences and the factors that affected the described outcomes.

Model Opening Question: As is said at the beginning we're interested in the way information is used by school administrators. To talk about this I want to identify 2 or 3 particular situations. I would like you to think back over the past two years and try to recall two or three <u>significant occurrences in the life of the program</u> here at _____ school. I realize that this question is somewhat vague, but it is vague on purpose. I want to get your impression of what was importnat rather than mine. Try and recall a few different <u>occurrences that you thought were significant</u> in determining the shape and character of the program during the last two years. For now I'd just like to list two or three such occurrences. We'll discuss the details later.

sub topics:

--changes (personnel, organization, goals, curriculum, materials, activities, attitudes, other milestones, etc.)

--rejected alternatives

--reinforcements in points of views, attitudes

Topic Area 4: <u>Factors Affecting the Specified Occurrences</u>

To determine the relative contribution of evaluation information in the total decision context, the respondents will be

asked to describe the factors that influenced their opinions and actions in the program events they just identified. Among the constituent influences in the situation might be such things as: the respondent's personal educational beliefs and predispositions, the respondent's first-hand observations, information from other school site personnel, information and suggestions from district staff, comments from parents and community members, contact with state and federal program offices, information from evaluations, fiscal pressures, etc.

Model Opening Question: I would like to discuss each of these occurrences in greater detail so I can try to understand the different factors that influenced peoples' actions and points of view. I want you to explain things to me in enough detail that I begin to see all the forces that were at work in the particular situation. Let's start with the case of the _____ _____. How did it happen that the school decided to _____? (Or, how did it happen that _____ _____occurred?)

sub topics:

--history; description of the occurrence

--different influences

--resolution process

Topic Area 5: The Role of Evaluation in the Identified Situations

No special attention was given to evaluation information in the previous stages of the interview. In many cases the respondents will have identified evaluation as one of the factors that influenced the actions they discussed. In the event that evaluation

was not mentioned, the question will be specifically asked by the interviewer at this stage in the interview.

Model Opening Question: Did evaluation make any difference in this situation?

Topic Area 6: The Role of Evaluation in General

To this point, evaluation has appeared as a secondary consideration in the interview. The situations identified by the respondents were allowed to define the scope of the discussion. Now, evaluation will be considered in its own right, and the respondent's wider knowledge and contact with evaluation will be investigated.

Model Opening Question: We've discussed _____ and _____ in great detail, and I think I understand the important factors involved in those occurrences (brief elaboration). Dr. Alkin and I are particularly interested in the usefulness of information from evaluations. I'd like to ask you to shift your thinking from these specific situations to thinking about evaluation in general. Will you take a minute to recall the program evaluations that have gone on in the past year or two; then, try to tell me what impact they had on you and on the programs at the school?

sub topics:

--level (within school, district sponsored activities, PQR and "mock review")

--characteristics of the evaluation (formal/informal, content, style, personalities, method of communication, etc.)

--its influence (on actions, attitudes, etc.)

--improving evaluation usefulness

Topic Area 5: The Role of Evaluation in the Identified Situation

After the more extensive discussion of evaluation it may be appropriate to repeat the earlier inquiry into significant occurrences. Certain subtle evaluation influences may have emerged from the lengthier discussion which were overlooked previously:

Topic Area 7: Additional Comments

At the conclusion of the interview, there will be a brief open-ended discussion period. Respondents will be given the opportunity to modify or expand their previous comments and clairfy any misinterpretations.

Model Opening Sentence: Before we conclude, I want to give you an opportunity to make any additional comments about our discussion. Is there anything you feel should be clarified or expanded with respect to the situations you identified, the various factors you singled out or about evaluation in general?

Thank you very much for your cooperation.

Revised 2/15/80

APPENDIX D

Interview Summary Form

Evaluation-User Survey

Interview Summary Form

Name of Interviewer: _____ School: _____

Respondent: _____

Title: _____

Step I. After completing the interview, but before listening to the recording:

1. Based on the complete interview, describe in one paragraph the specially-funded programs operating that this school.

_____

_____

_____

_____

_____

_____

_____

_____

_____

2. In one paragraph, describe the respondent's duties and responsibilities, particularly as they involve the special programs you discussed.

_____

_____

_____

_____

_____

_____

_____

_____

_____

3. In one paragraph <u>each</u>, describe the significant occurrences identified by the respondent and discussed in the interview.

Situation 1:

Situation 2:

4. For each occurrence: A. List (in approximate order of importance) the factors that influenced the final outcome.
B. Summarize in one paragraph the interrelationships among these factors.

Situation 1:

Situation 2:

5. Was evaluation information a factor in each of these situations? For each occurrence summarize in one paragraph the role of evaluation.

Situation 1: _____

Situation 2: _____

6. In one paragraph each describe the respondents experiences with evaluation at the "school", "district", and "state" levels. Indicate a) type of evaluation, b) its influence/ usefulness, c) factors contributing to its influence, d) ways of improving evaluation.

"school level" activitities: _____

"district level" activities: _____

PQR & mock review: _____

7. Summarize in one paragraph any additional comments that were important.

In one paragraph describe the interview context--the salient features of the setting, the participants and the interaction.

Step II: Replay the interview tape. (Set the counter at zero 000 at the beginning of each new side.)

As you listen to the interview:

1. Make additions/corrections to the descriptive paragraphs you wrote in Step I.
2. Select important quotes to illustrate key features of the interview.
3. Write out the quotes on the following pages.
   a. First indicate in a sentence or two what is being discussed immediately prior to the quote, i.e., some context for the remark. If it is an answer to a particular question, give the question.
   b. Write the quote as accurately as you can.
   c. Don't forget to indicate the tape counter reading at the beginning and end of each quote.

Key quotes:

Topic Area _____ Tape Counter at beginning of quote _____ Side A B ?

Context/Question: _____

_____

_____

Quote: _____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Tape Counter at end of quote _____ Side A B ?

Topic Area _____ Tape Counter at beginning of quote _____ Side A B ?

Context/Question: _____

_____

_____

Quote: _____

_____

_____

_____

_____

_____

_____

_____

Tape Counter at end of quote _____ Side A B ?

Topic Area _____ Tape Counter at beginning of quote _____ Side A B ?

Context/Question: _____

_____

_____

Quote: _____

_____

_____

_____

_____

_____

_____

_____

Tape Counter at end of quote _____ Side A B ?

APPENDIX E

Coding Form

SURVEY INSTRUMENT FOR USER INTERVIEW TAPE LISTENING

CSE Evaluation Use Project

School/Respondent Code _____ Reviewer(s) _____/_____

Significant Occurrence # _____

1. What action was taken in the "significant occurrence"?

(Check one)

_____ applying/not applying for funding
_____ developing or implementing a program after receiving new or increased funding
_____ modifying a program after a decrease or termination of funding
_____ responding to non-monetary changes in the integration plan
_____ modifying classroom organizational patterns (e.g. teaming, groupings, etc.)
_____ modifying staff person responsibilities
_____ switching, adding, or deleting instructional materials, texts, management systems, etc.
_____ implementing or modifying instructional guidelines or objectives across classrooms IN RESPONSE TO DISTRICT MANDATE
_____ implementing or modifying instructional guidelines or objectives across classrooms VOLUNTARILY
_____ initiating or modifying staff development activities
_____ hiring new teachers
_____ hiring new auxiliary staff
_____ hiring aides/paraprofessionals
_____ initiating or modifying parent activities
_____ changing school schedules
_____ modifying the school decisionmaking or governance structure
_____ initiating or modifying activities related to student behavior or discipline
_____ initiating or modifying activities related to evaluation or needs assessment
_____ modifying the physical plant
_____ other

In a brief phrase describe the "significant occurrence":

_____

What was the specific decision under consideration?

_____

_____

I. Initiating Events.

2. Was the initial prompt to school action a result of:

_____ an increase in funds for an existing specially funded program? (Identify _____)

_____ a decrease or halt in funds for an existing specially funded program? (Identify _____)

_____ initiation of a new specially funded program in the school? (Identify the program _____)

_____ some other action/request by the State Department of Education?

_____ some other action/request by the school district administration?

_____ parent action

_____ mandatory reaction to demographic changes

_____ voluntary reaction to demographic changes

_____ needs assessment data collected for the Consolidated Application

_____ voluntarily-collected needs assessment data

_____ suggestion/comment from within the school

_____ Other origin (Specify_____)

_____ Don't know

3A. If the prompt was external, to whom was the prompting information first communicated?

_____ principal
_____ asst. principal
_____ staff person(s)
_____ teacher(s)
_____ aide(s)
_____ other (Identify _____)
_____ don't know

3b. If the prompt was internal to the school:
Was the initial idea that a change was needed attributed strongly to one particular person or organizational group?
_____no
_____yes

If Yes, identify the person or group:

_____ the principal
_____ an asst. principal
_____ a staff person
_____ a teacher
_____ a teacher group
_____ a mixed teacher/other staff group
_____ parent(s)

Person/group includes respondent?

_____ no
_____ yes
_____ don't know

4. Did the prompt carry with it a suggested course of action?
_____ no action concurrently suggested
_____ one specific action suggested
_____ more than one action option suggested
_____ unknown/not applicable

5. In what context was the prompt first discussed within the school?
_____ privately, among a few persons
_____ in a public setting, but not at a formal meeting (e.g. in the lunchroom)
_____ in a formal setting (e.g. staff mtg.)
_____ don't know

6. Sometimes people create a strategy or set of steps for coming to a decision about what to do, i.e. they establish a procedure for planning and choosing their action. Who, if anyone, was most responsible for determining such a procedure in this case?

_____ Not applicable, no procedural plan established
_____ don't know
_____ the principal
_____ a staff person(s)
_____ a teacher
_____ a group of people, collectively

II. Charting the "Decision Process":

Person codes are:

1	Principal
2	Staff person
3	Individual teacher
4	Individual aide
5	Individual parent
6	"Executive committee," "Leadership committee," etc.
7	Representative (i.e. elected or appointed) teacher committee
8	Full staff
9	Informal teacher group
10	"School Site Council" (a formal teacher and parent group)
11	Formal parent committee
12	All parents of students at the school
13	Informal parent group
14	Informal aide group
15	Informal mixed group of school professionals only
16	Informal mixed group including professionals, semi-professionals, and parents
17	Area staff personnel
18	Research & Evaluation Office personnel
19	Consultants
20	Auxiliary personnel
21	Other (Identify _____)

Information codes are:

A	Demographic data gathered for Consolidated Application or other specific purpose
B	Consolidated Application Plans and Guidelines
C	Other program guidelines
D	Comprehensive schoolwide needs assessment data (mandatory)
E	Smaller scale needs assessments or survey (voluntary)
F	CTBS scores
G	SES scores
H	Classroom tests
I	Test scores, undifferentiated
J	Area staff input (advice, suggestions, recommendations)
K	Research & Evaluation input
L	Collegial advice from principals at other schools
M	Collegial advice from staff at other schools
N	Principal's beliefs and opinions
O	Staff persons' beliefs and opinions
P	Teachers' beliefs and opinions
Q	Principal's observations (including informal data collection)
R	Staff persons' observations
S	Teachers' observations
T	Parent input
U	Aide input
V	Information from educational research or professional publications
W	Information from other media (publications, mags, TV, etc.)
X	Other information sources
Y	Unclear information sources
Z	PQR
AA	Publishers representatives or materials
BB	Budget constraints

Code the sequence of interactions and information use which began with recognition of the desirability or need for action and which eventuated in the school action described to us. Create a separate person-information string for each discrete meeting (or activity between meetings) that was described. A person-information string is a single person or group code followed by as many information source codes as apply.

7. Code the events relating to recognition/identification of the prompt. (Stage I)

8. Code the events involved in conducting the decision process, up to and including the choice of the final action (embodied in the "significant occurrence"). (Stage II)

9. Code the events involved in "signing off" or ratifying the final choice. (Stage III)

10. Code the audiences to whom the action plan was diffused or disseminated.

11. How much data do you, the coder, feel was missing in the preceeding account of the decision process?
    _____ very little
    _____ a moderate amount
    _____ a great amount

12. How confident are you, the coder, in the accuracy of the sequencing of events as coded?
    _____ not very confident
    _____ moderately confident
    _____ very confident

13. How much time elapsed between the initial prompt to action and the selection of the final plan?

_____ 3 days or less
_____ 4 days to 2 weeks
_____ 2 weeks to 1 month
_____ over 1 month
_____ don't know

14. Including the final plan, how many options were considered?

_____ 1 only
_____ 2
_____ 3
_____ 4
_____ don't know

16. Was there an <u>express effort</u> to search out information relevant to the school's choice of action? (Check all that apply.)

_____ don't know/systematic search not apparent
_____ polling or surveying persons
_____ <u>rescrutinizing</u> "old" available data (e.g. from needs assessments or testing). Note: This does not mean just citing such data as having been used.
_____ conducting a literature search
_____ contacting expert sources (includes district staff)
_____ delegating fact-finding responsibility to someone
_____ other (Specify_____)

17. Was the "final" action plan tried out and then modified on the basis of experience in the school?

_____ yes
_____ no
_____ don't know
_____ not applicable (no action, future action, etc.)

If yes, whose reactions/opinions were cited as most important in reassessing the plan? (Check all that apply.)

_____ principal
_____ staff person
_____ teachers
_____ aides
_____ parents
_____ students

18. Was a procedure established for monitoring the implementation and/or outcomes of the plan of action?

_____ no
_____ yes
_____ don't know
_____ not applicable (no action, action in future, etc.)

If yes, who was primarily responsible for the task?

_____ principal
_____ staff person
_____ teacher
_____ committee
(Give composition _____)
_____ unclear

19. Did the respondent provide other important pieces of information relating to the significant occurence that were not asked for in any of the previous items?

_____ yes
_____ no
If so, please describe:_____

_____

_____

_____

APPENDIX F

Frequency of Information Use:
Comparisons Between Each Type of Significant
Occurrence and the Total Sample

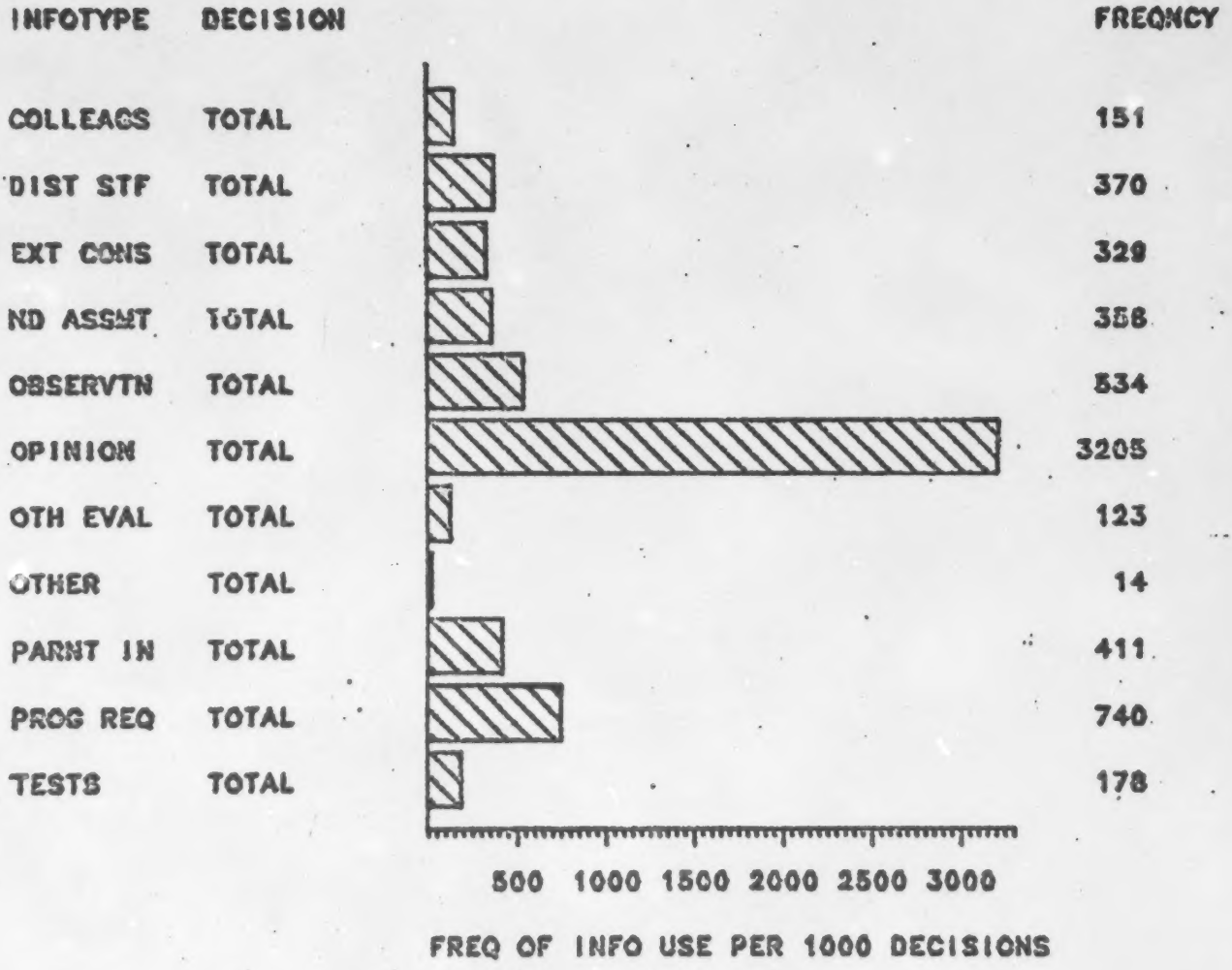

Figure E-1: Information Use, All Decisions (N=73)

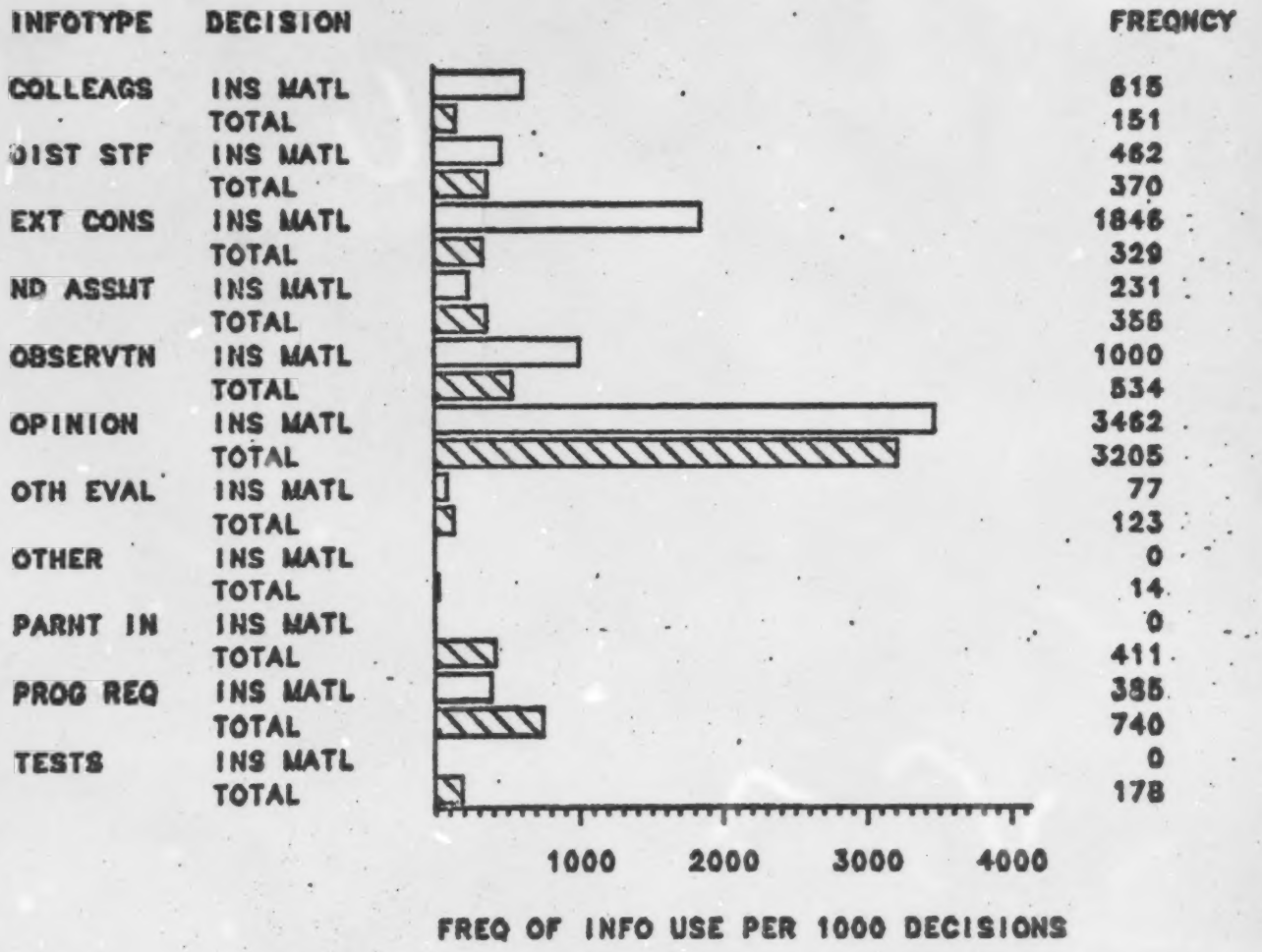

Figure E-2: Information Use, INS MATL(N=13) vs. TOTAL(N=73)

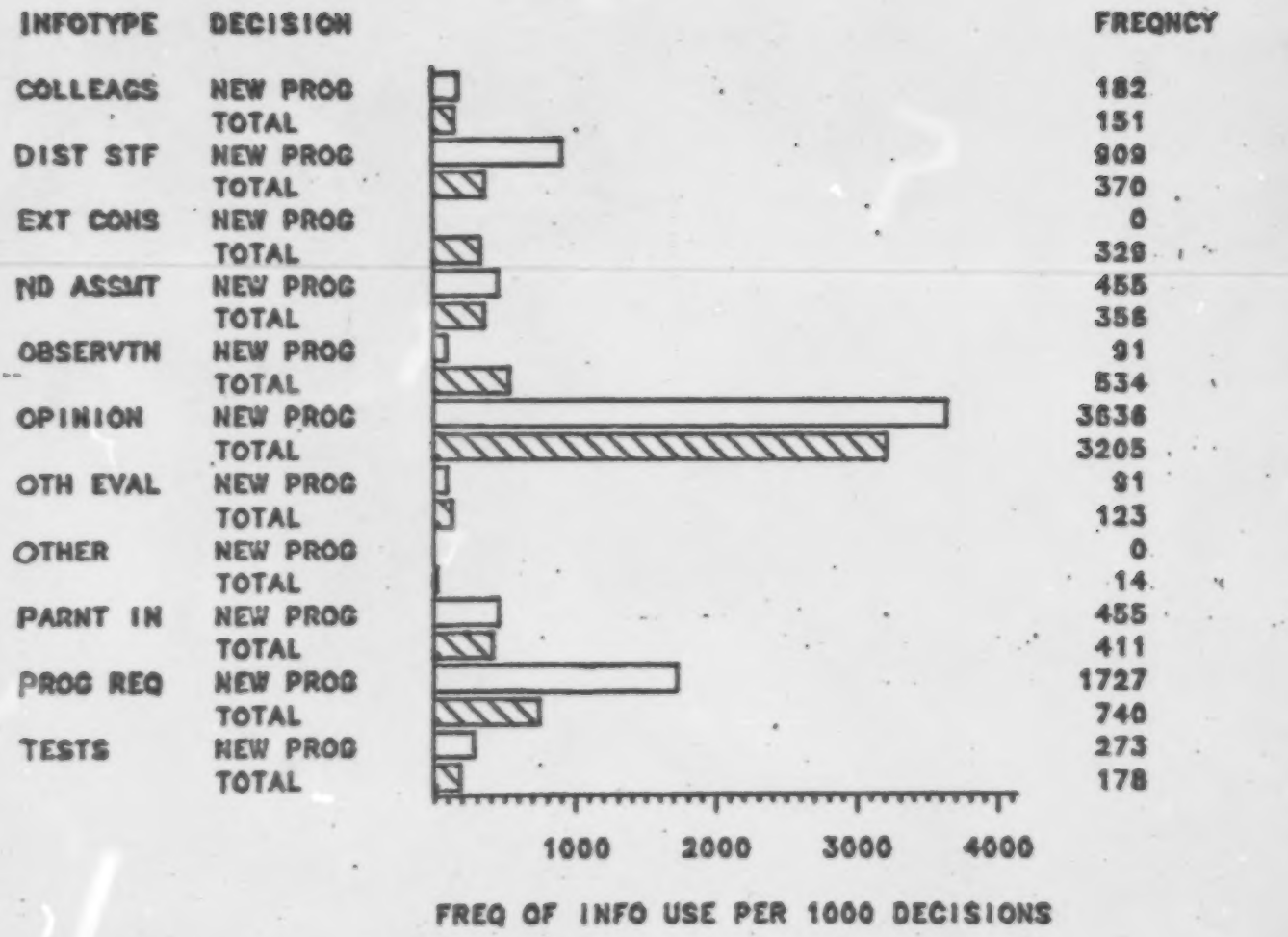

Figure E-3: Information Use, NEW PROG(N=11) vs. TOTAL(N=73)

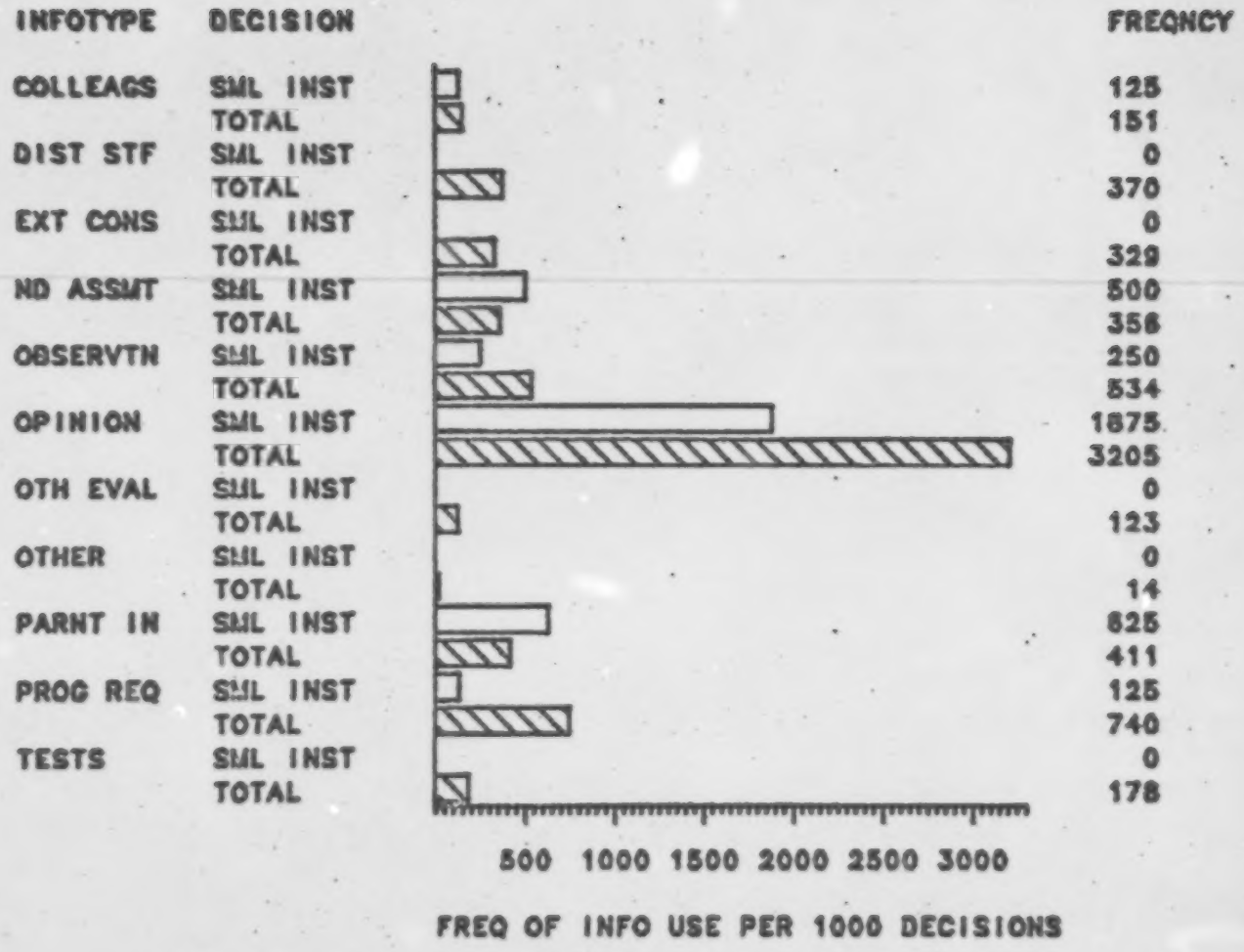

Figure E-4: Information Use, SML INST(N=8) vs. TOTAL(N=73)

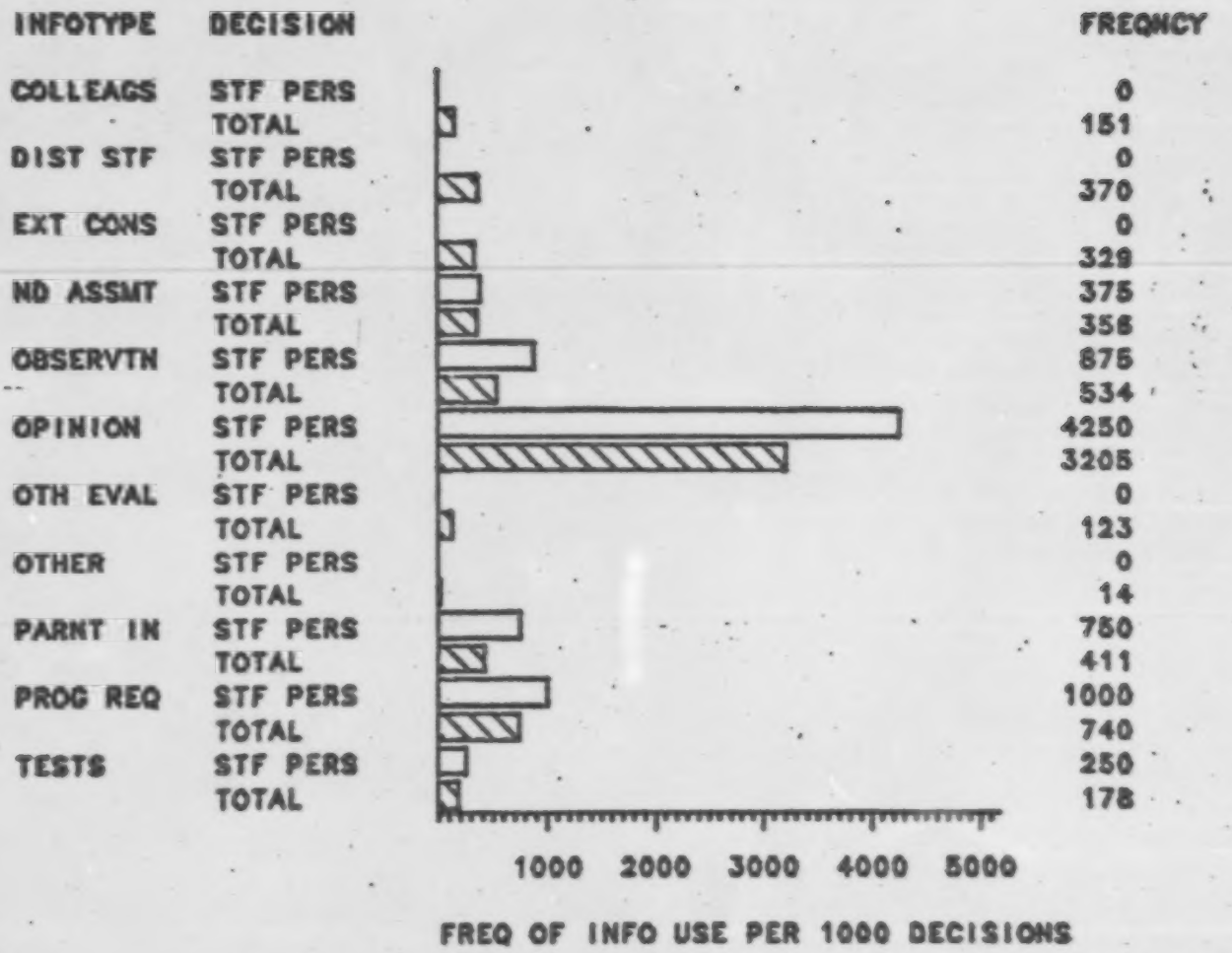

Figure E-5: Information Use, STF PER(N-8) vs. TOTAL(N=73)

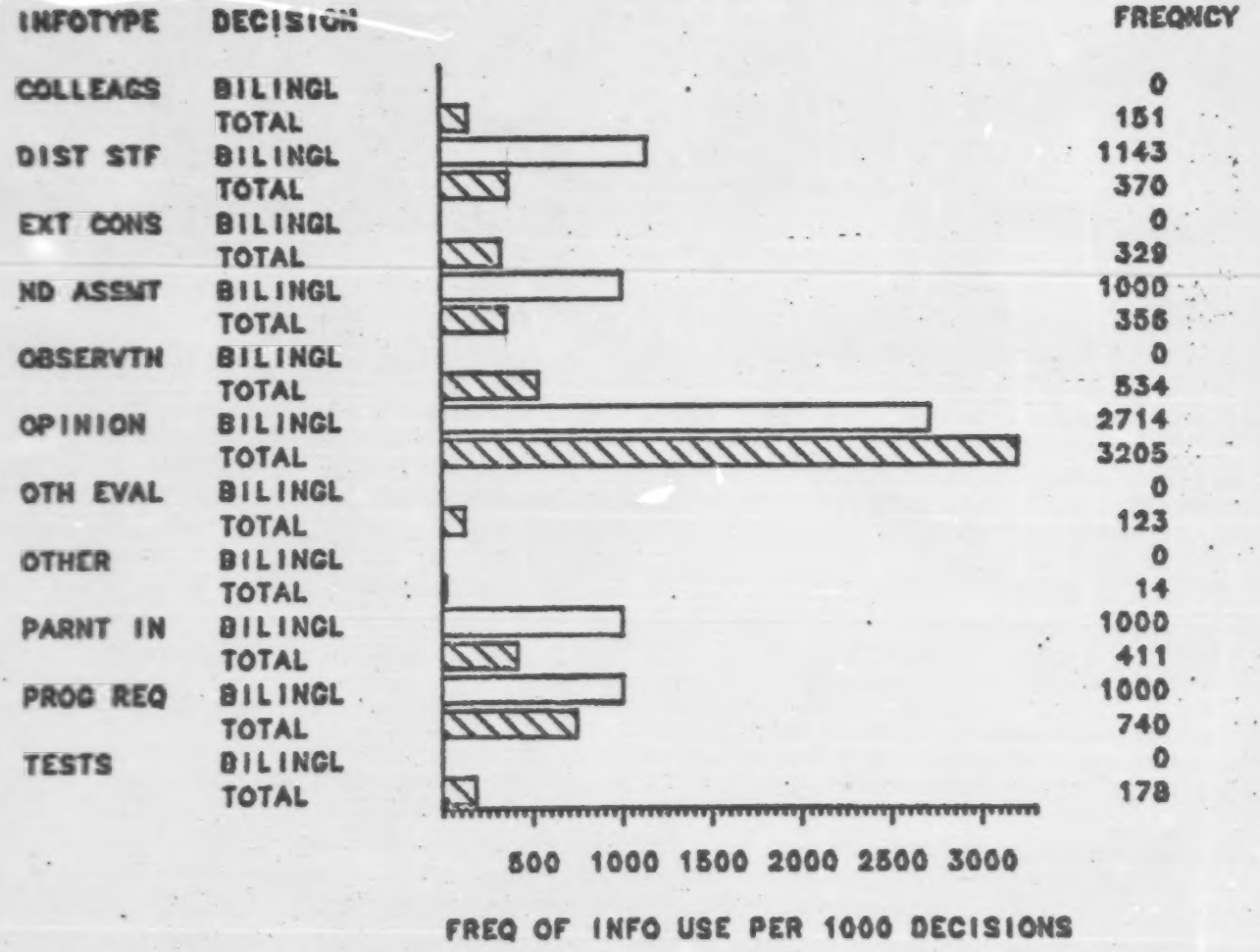

Figure E-6: Information Use, BILINGL(N=7) vs. TOTAL(N=73)

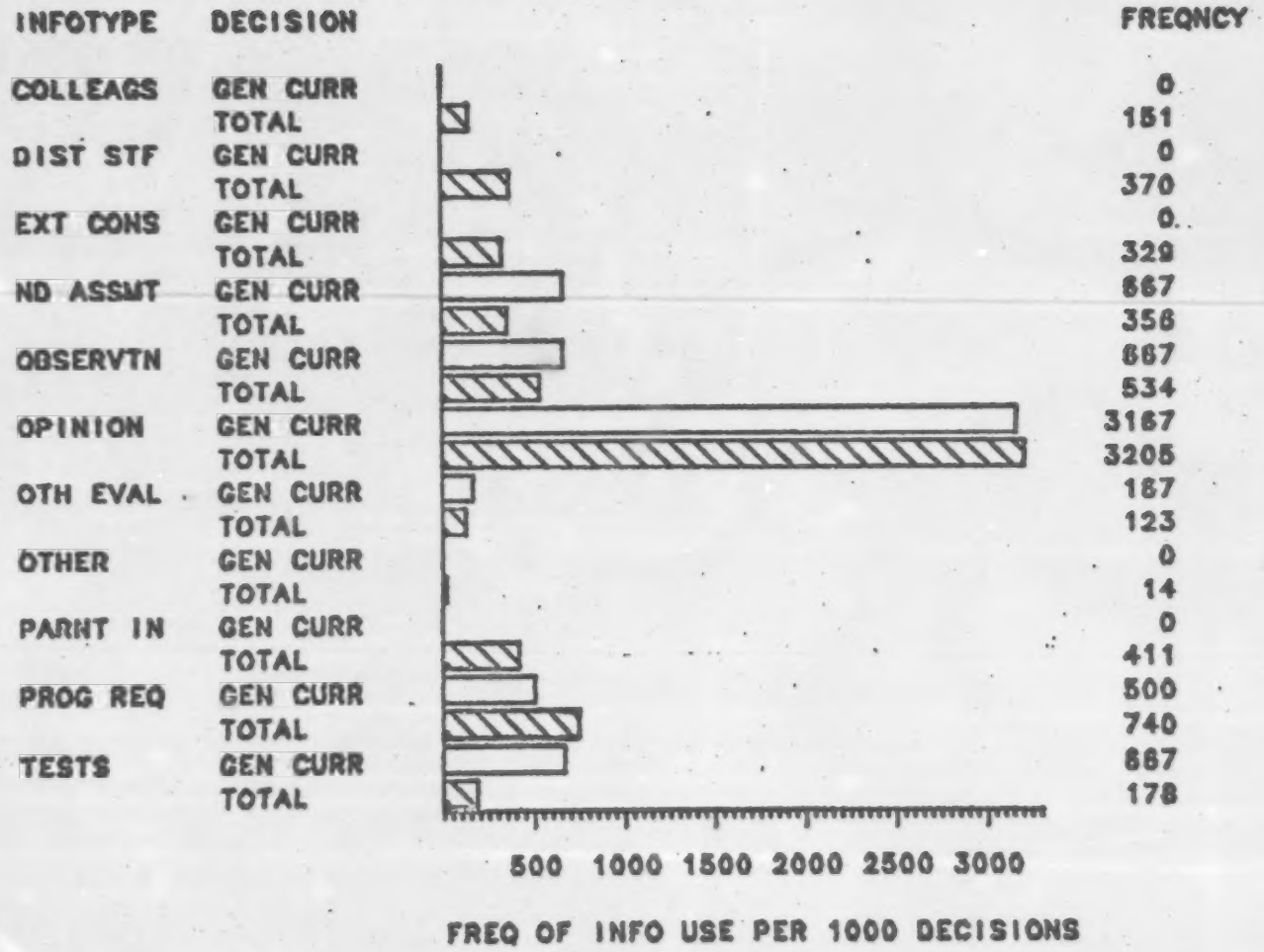

Figure E-7: Information Use, GEN CURR(N=6) vs. TOTAL(N=73)

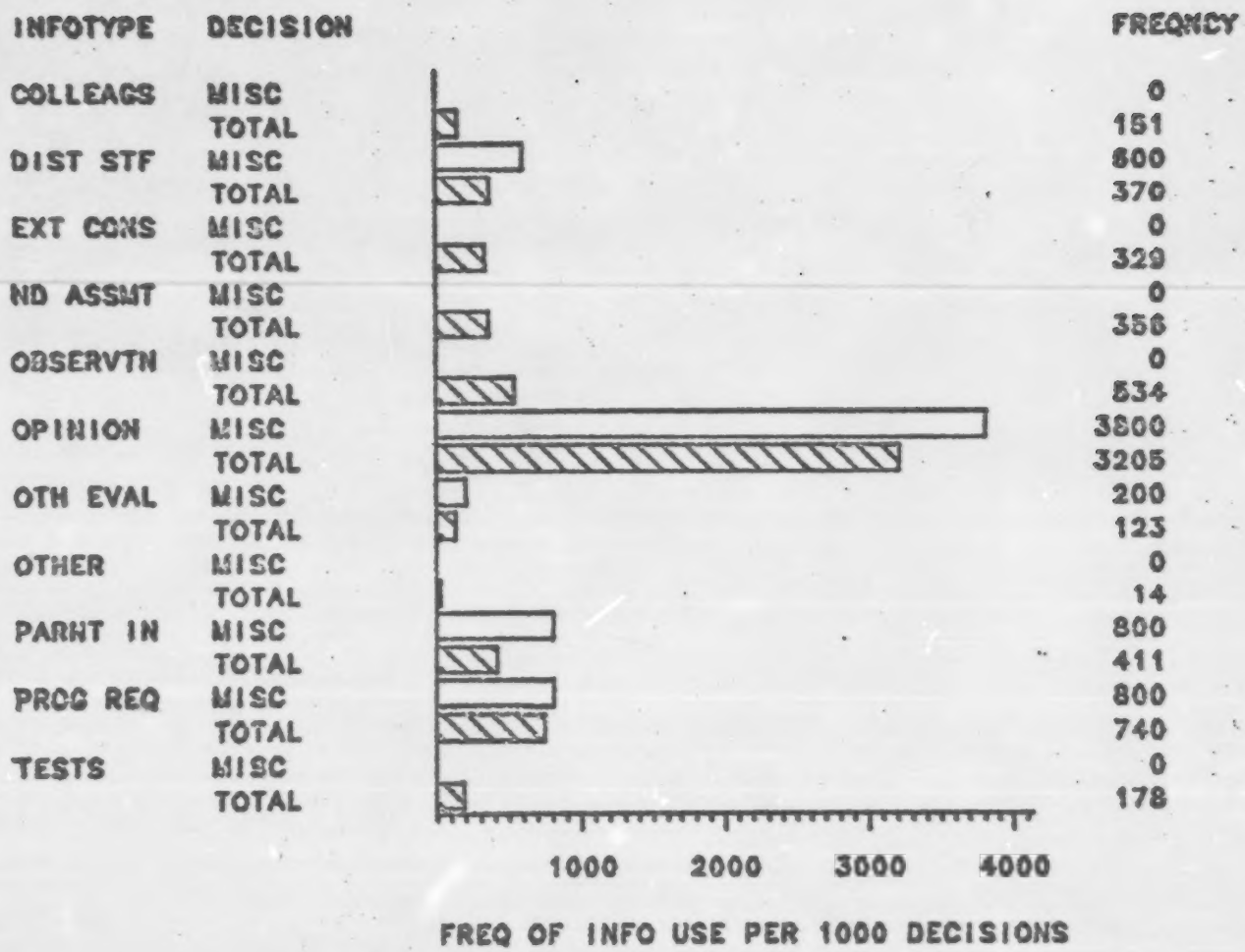

Figure E-8: Information Use, MISC(N=5) vs. TOTAL(N=73)

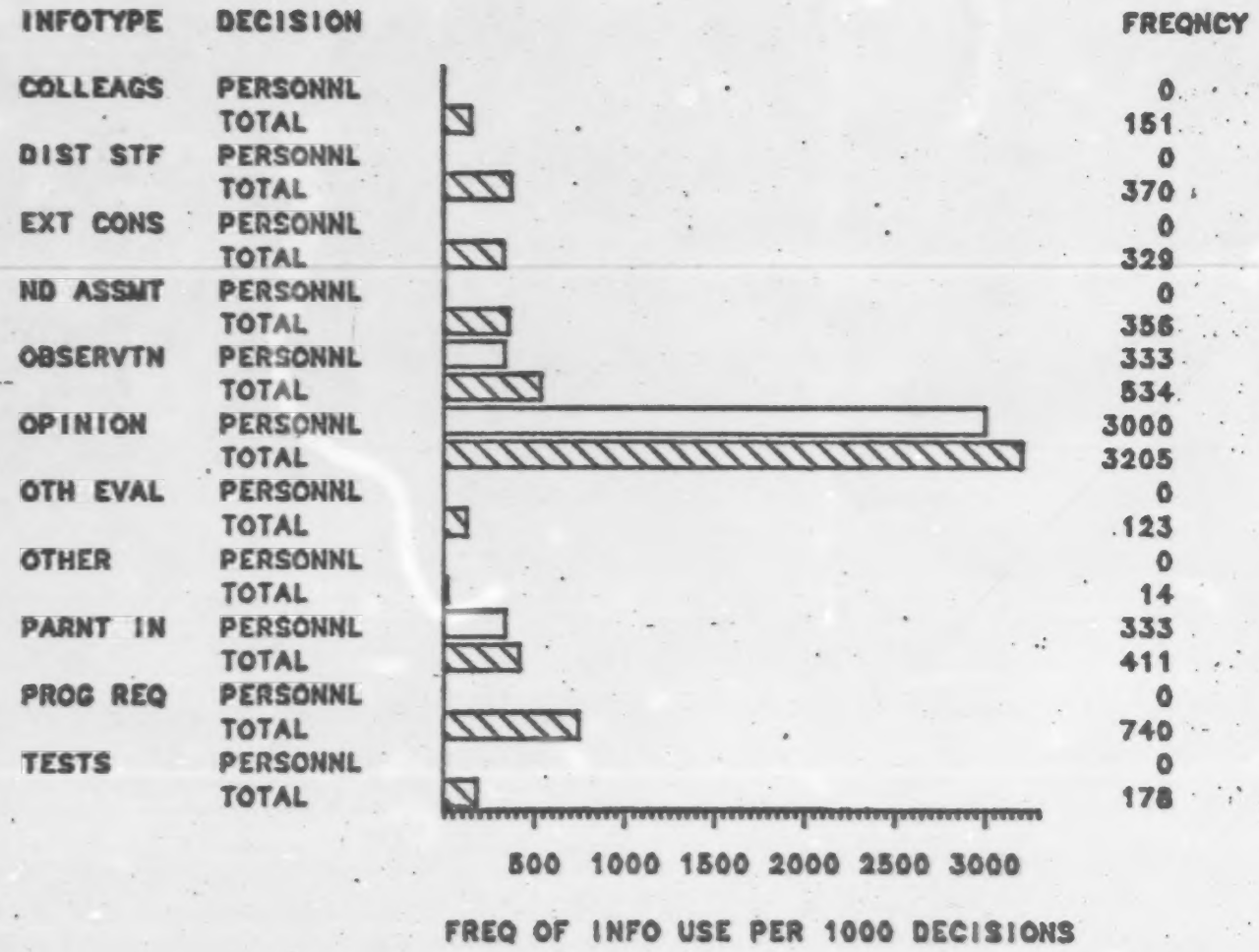

Figure E-9: Information Use, PERSONNL(N=3) vs. TOTAL(N=73)

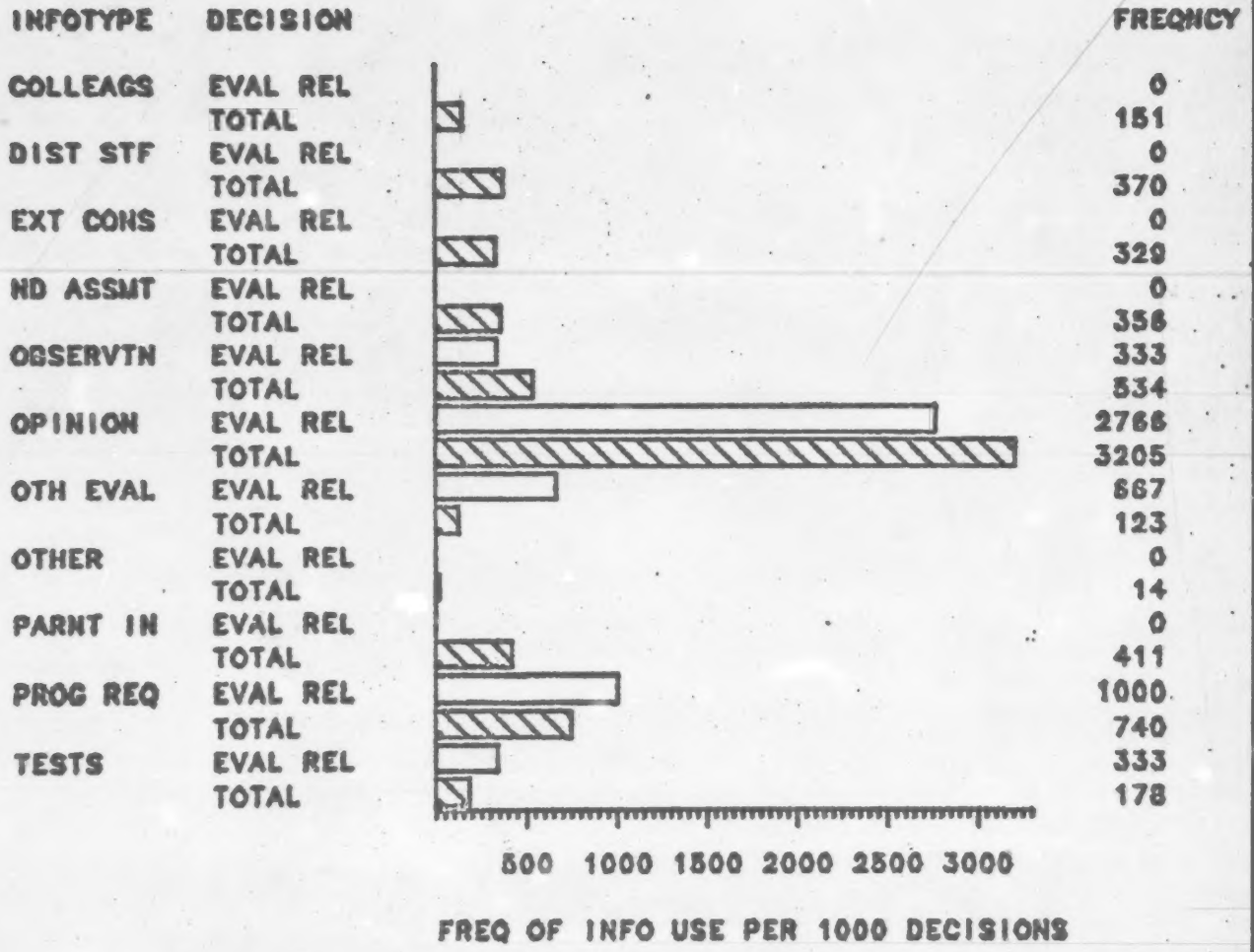

Figure E-10: Information Use, EVAL (N=3) vs. TOTAL(N=73)

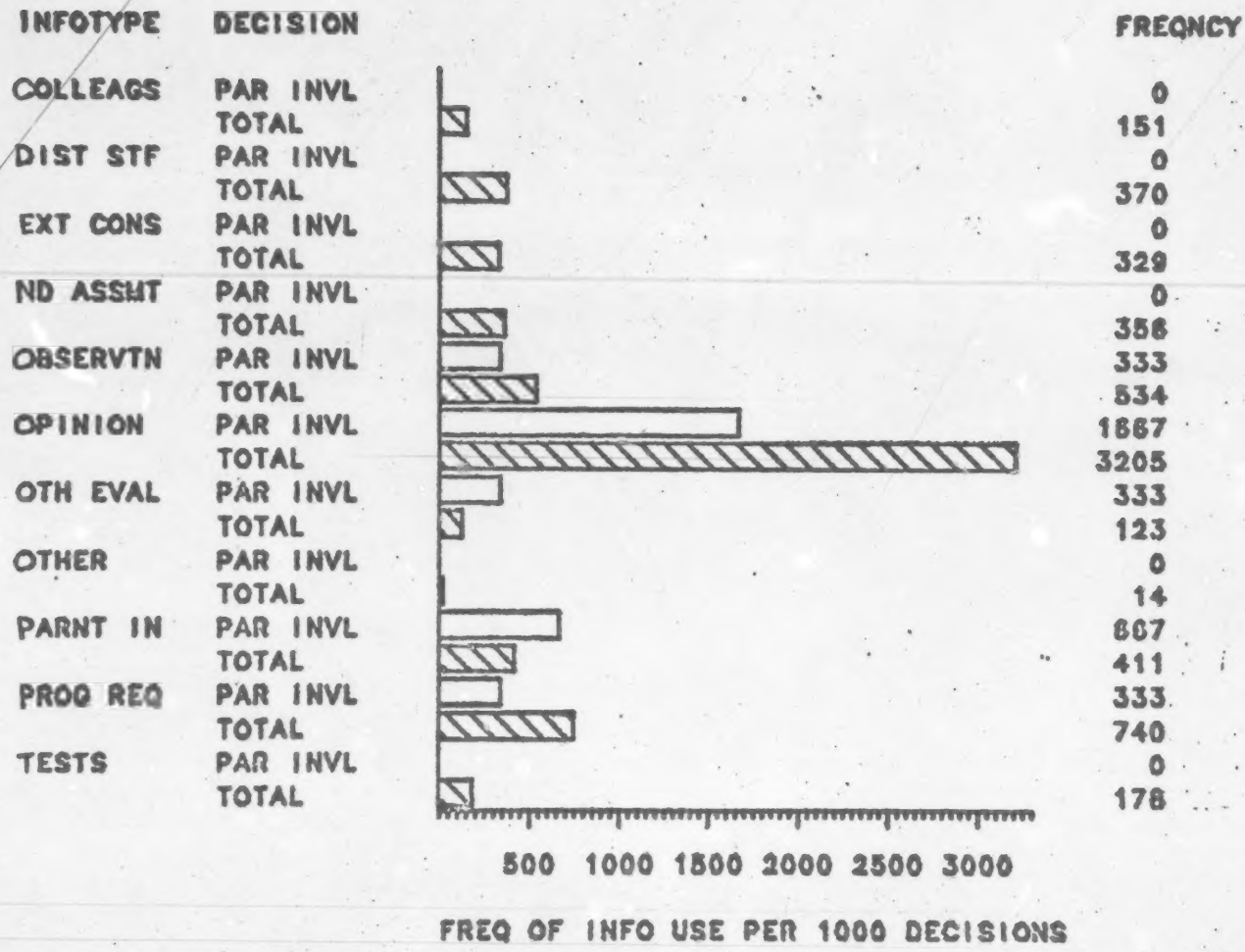

Figure E-11: Information Use, PAR INVL(N=3) vs. TOTAL(N=73)

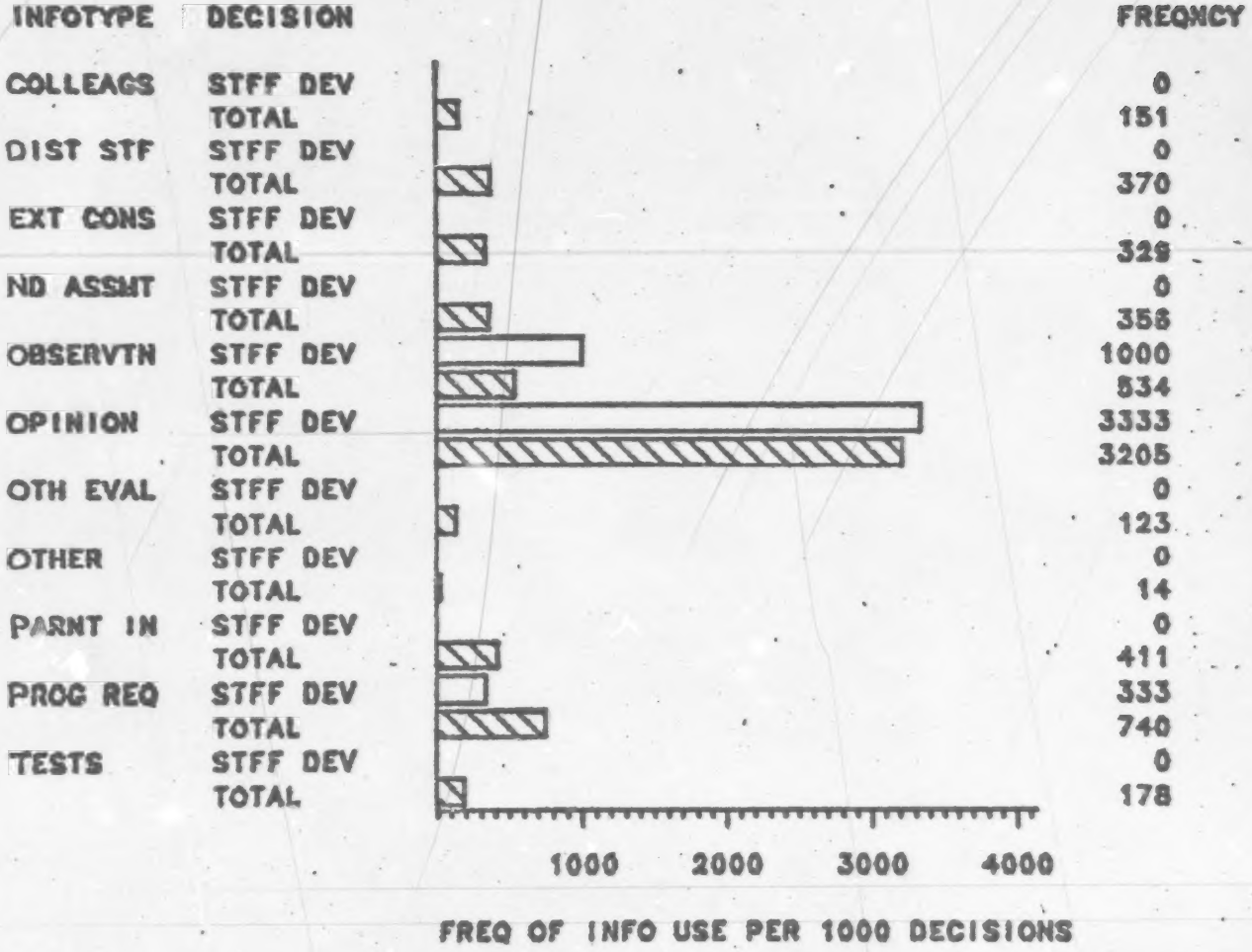

Figure E-12: Information Use, STFF DEV(N=3) vs. TOTAL(N=73)

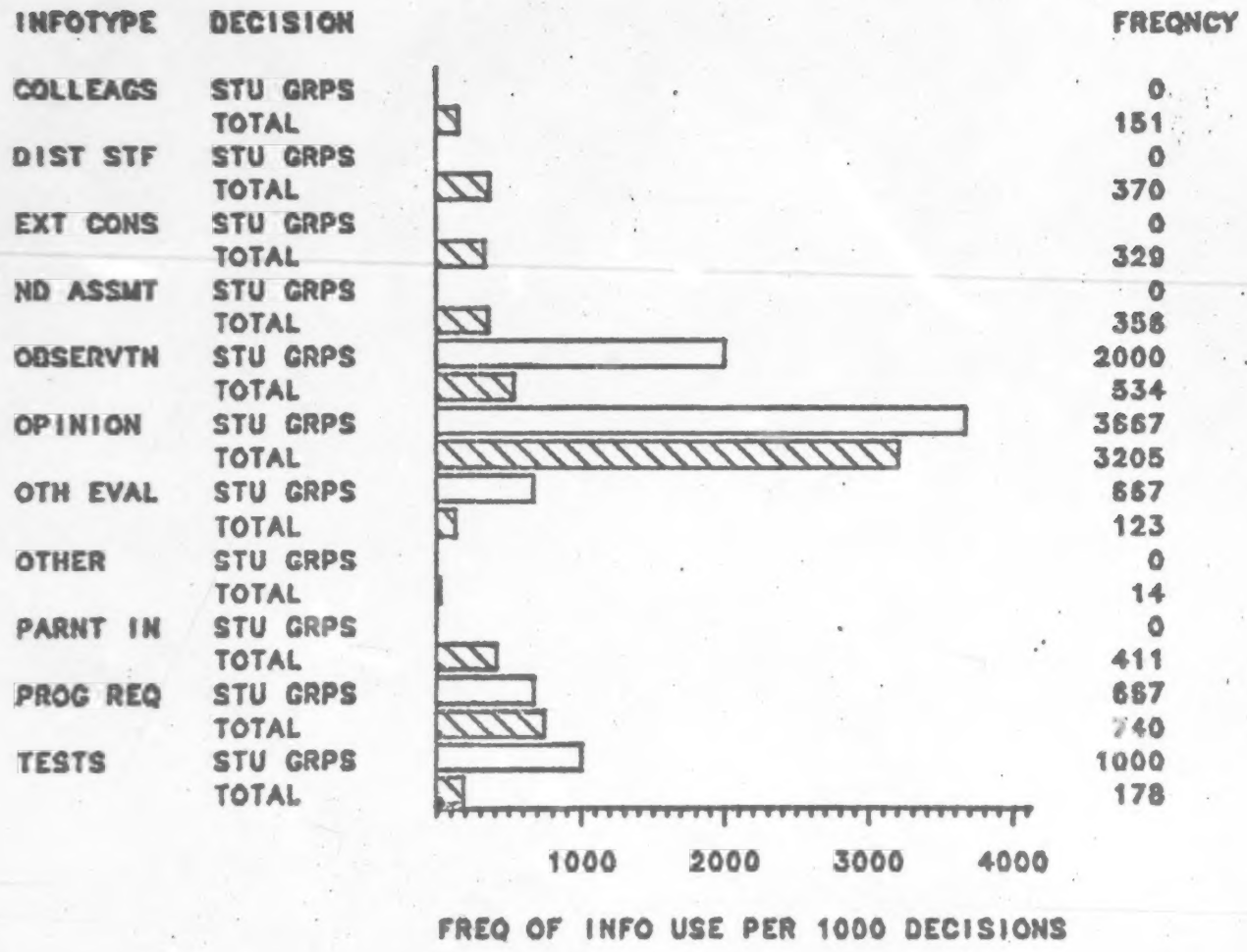

Figure E-13: Information Use, STU GRPS(N=3) vs. TOTAL(N=73)

APPENDIX G

Frequency of Information Use:

Comparisons Between Different Personnel Categories

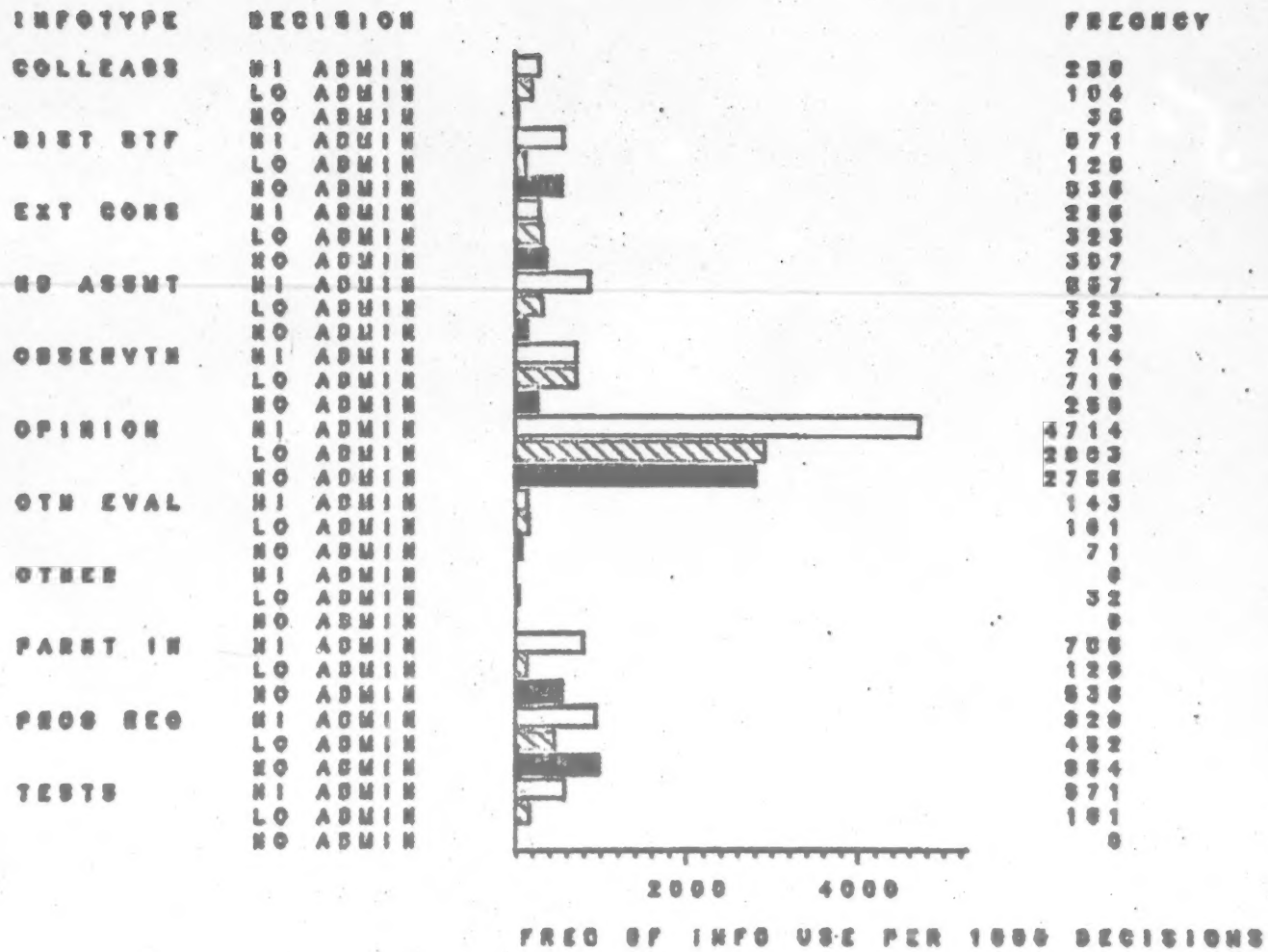

Figure F-1: Info Use, HI ADMIN(N=14) vs. LO ADMIN(N=31) vs. NO ADMIN(N=28)

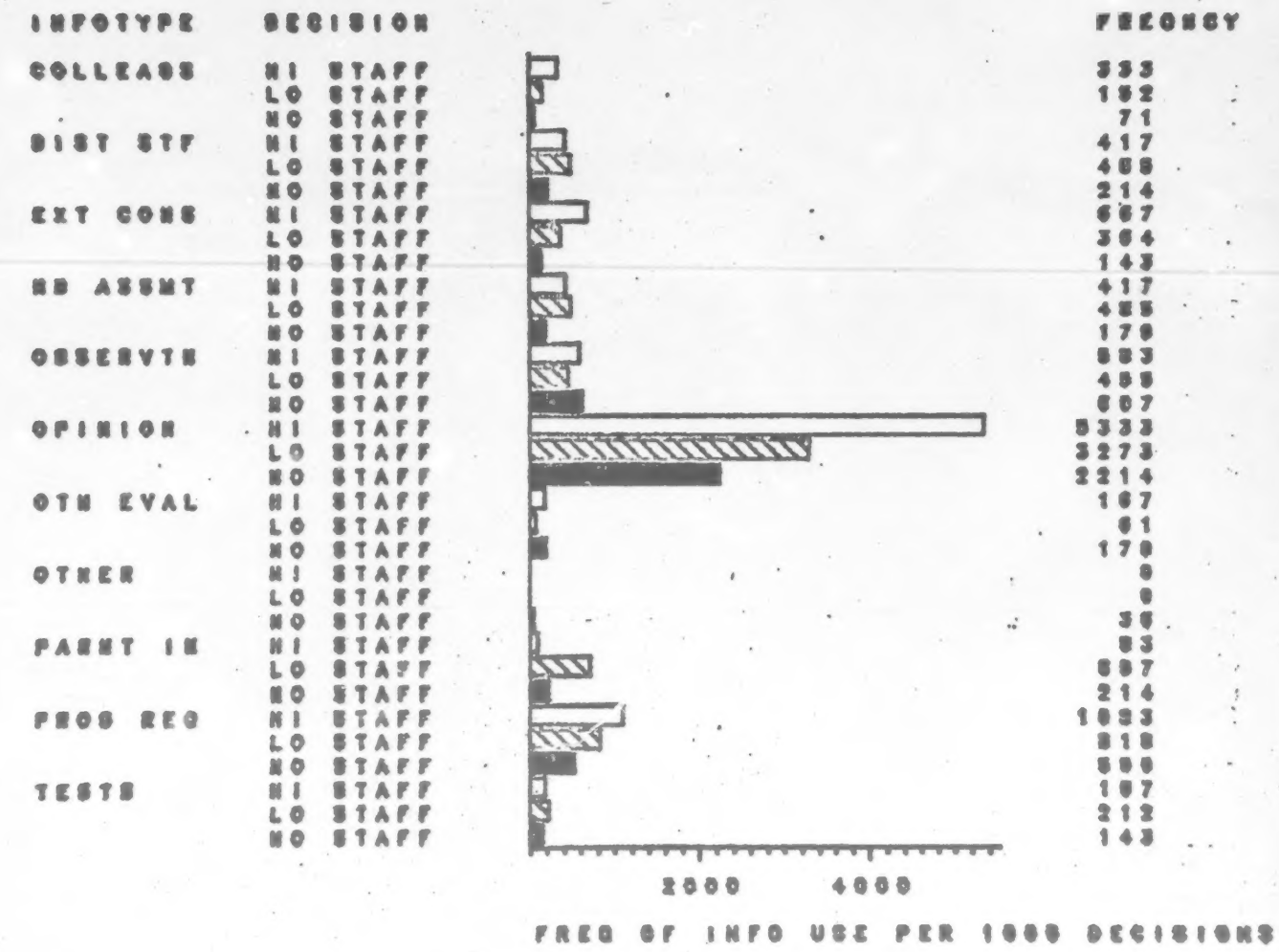

Figure F-2: Info Use, HI STAFF(N=12) vs. LO STAFF(N=33) vs. NO STAFF(N=28)

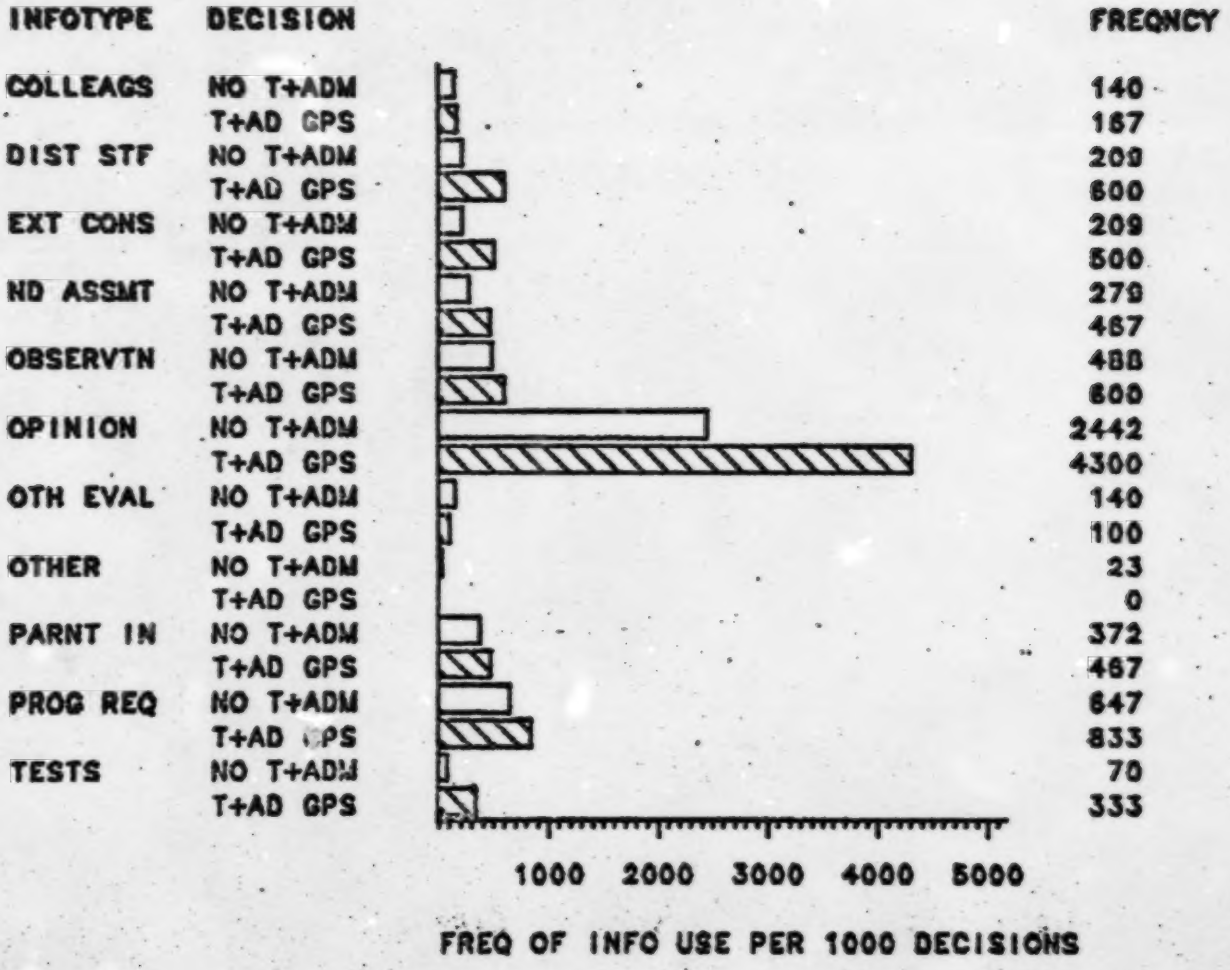

Figure F-3: Information Use, T+AD GPS(N=30) vs. NO T+ADM(N=43)

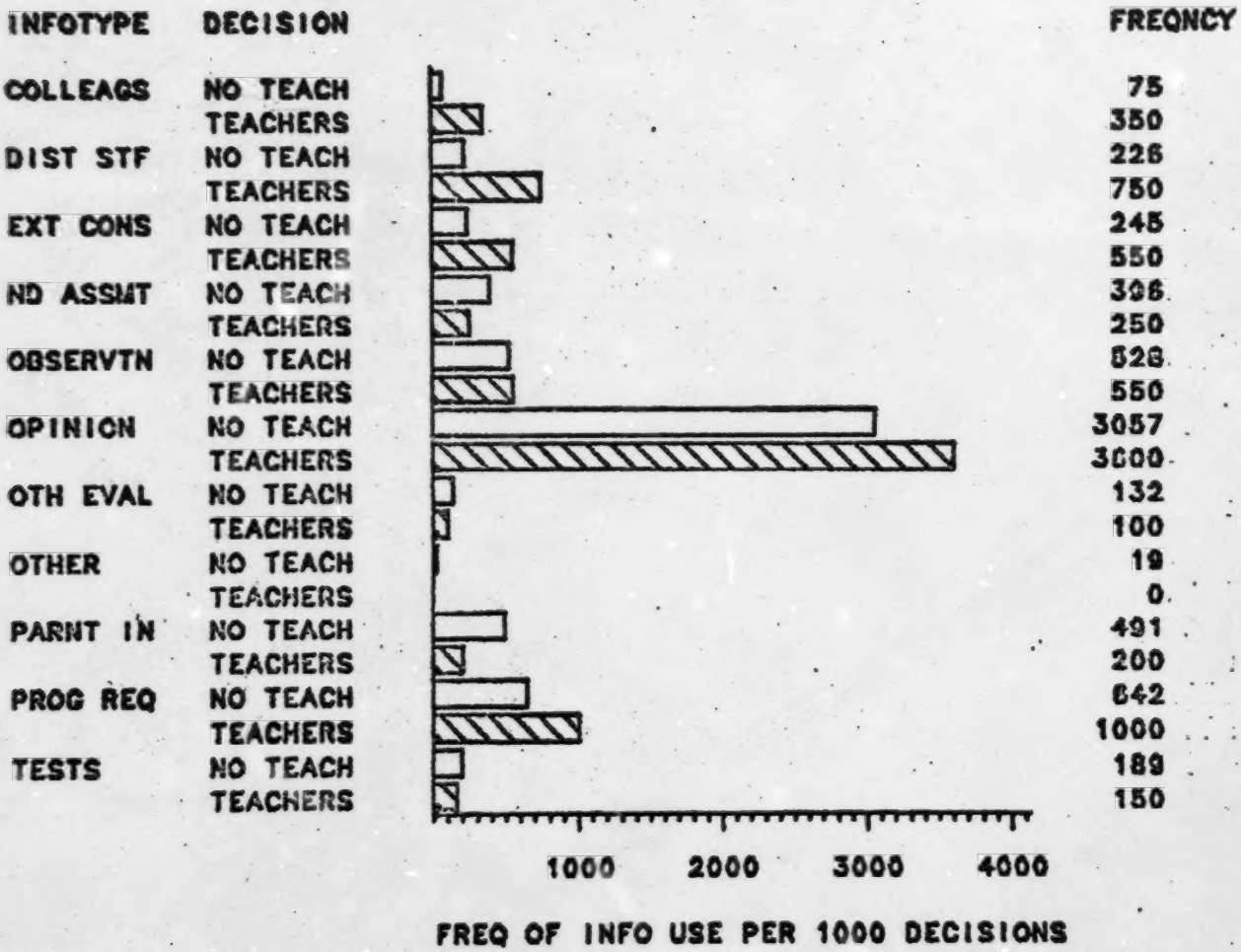

Figure F-4: Information Use, NO TEACH(N=53) vs. TEACH(N=20)

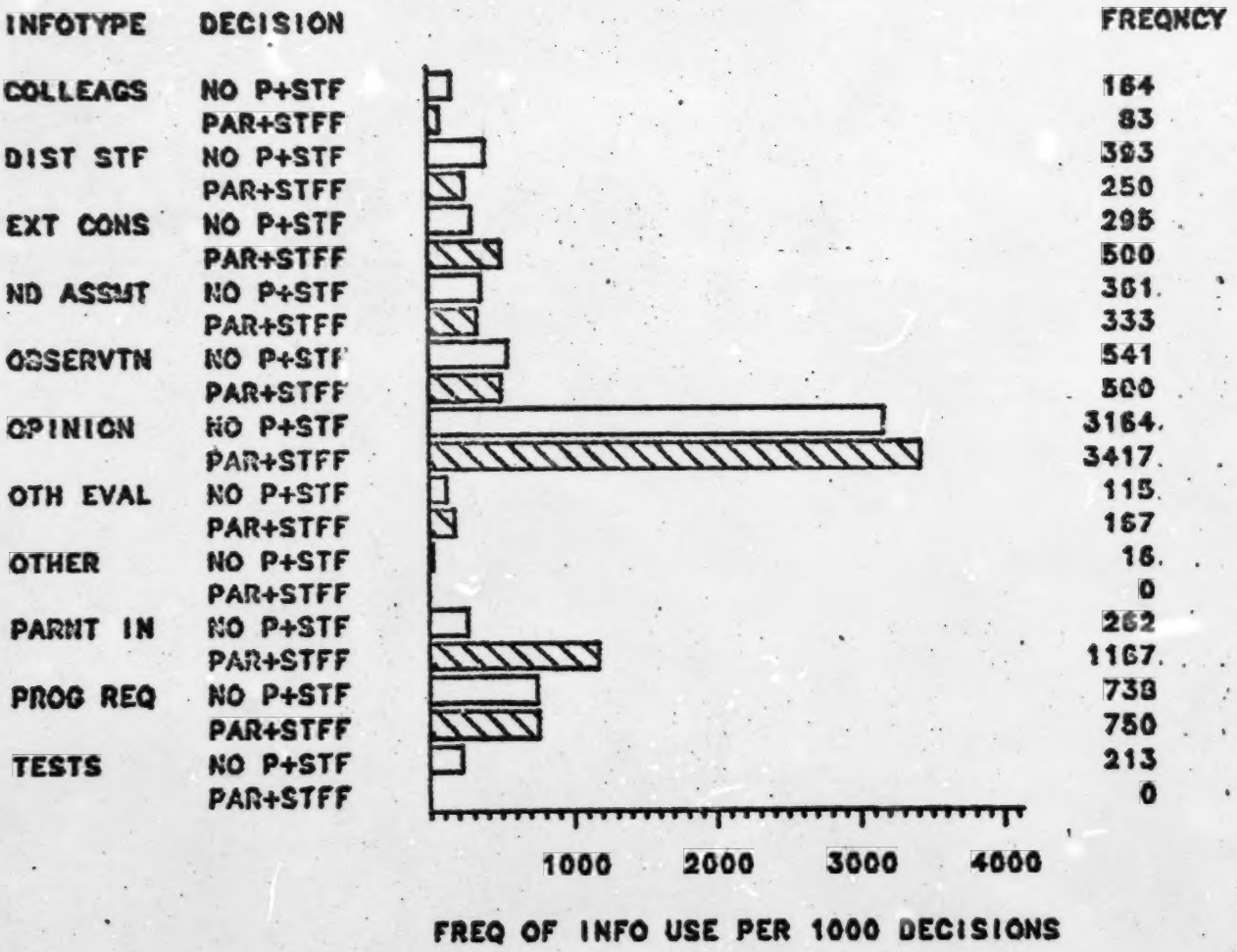

Figure F-5: Information Use, PAR+STFF(N=12) vs. NO P+STF(N=61)

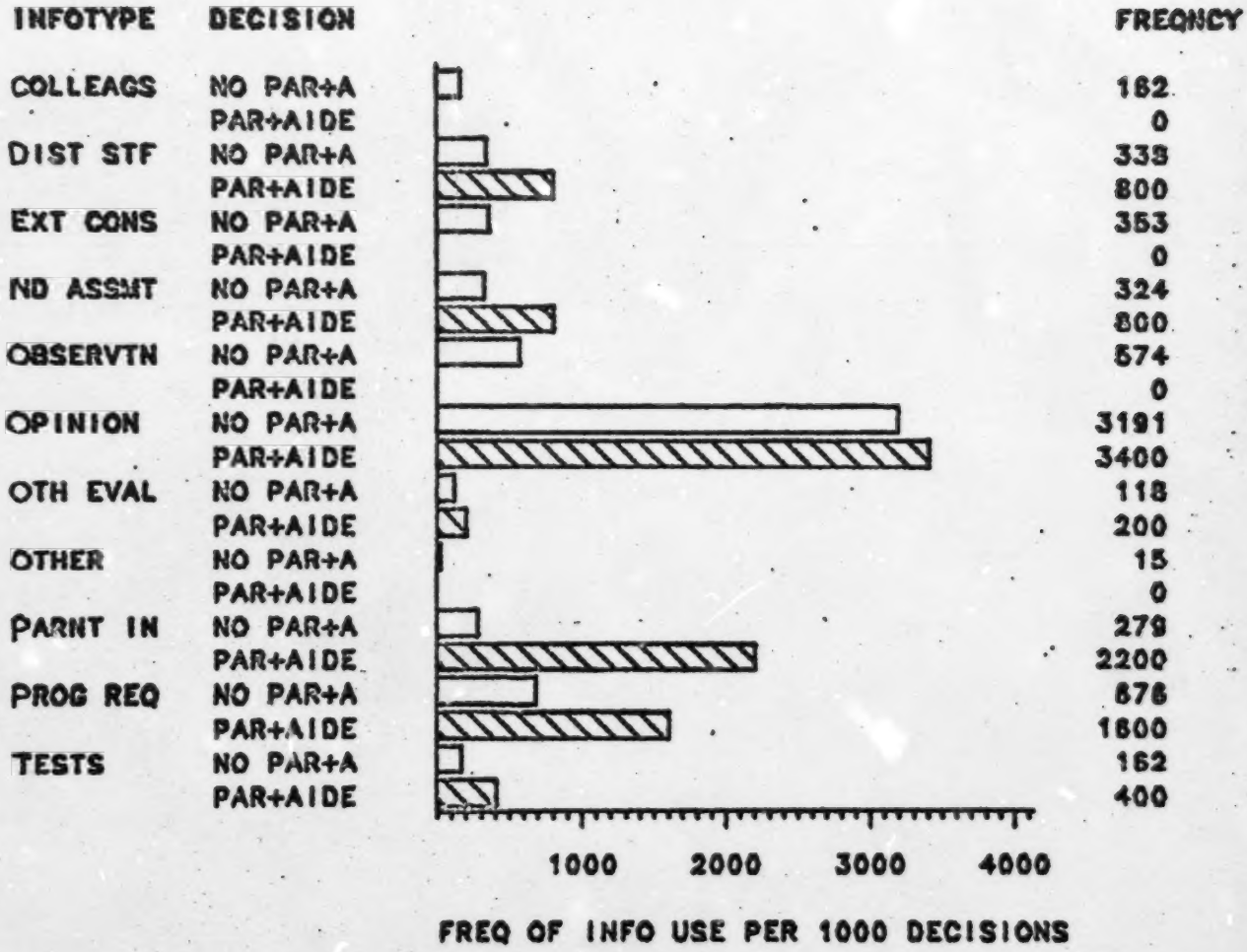

Figure F-6: Information Use, PAR+AIDE(N=5) vs. NO PAR+A(N=68)

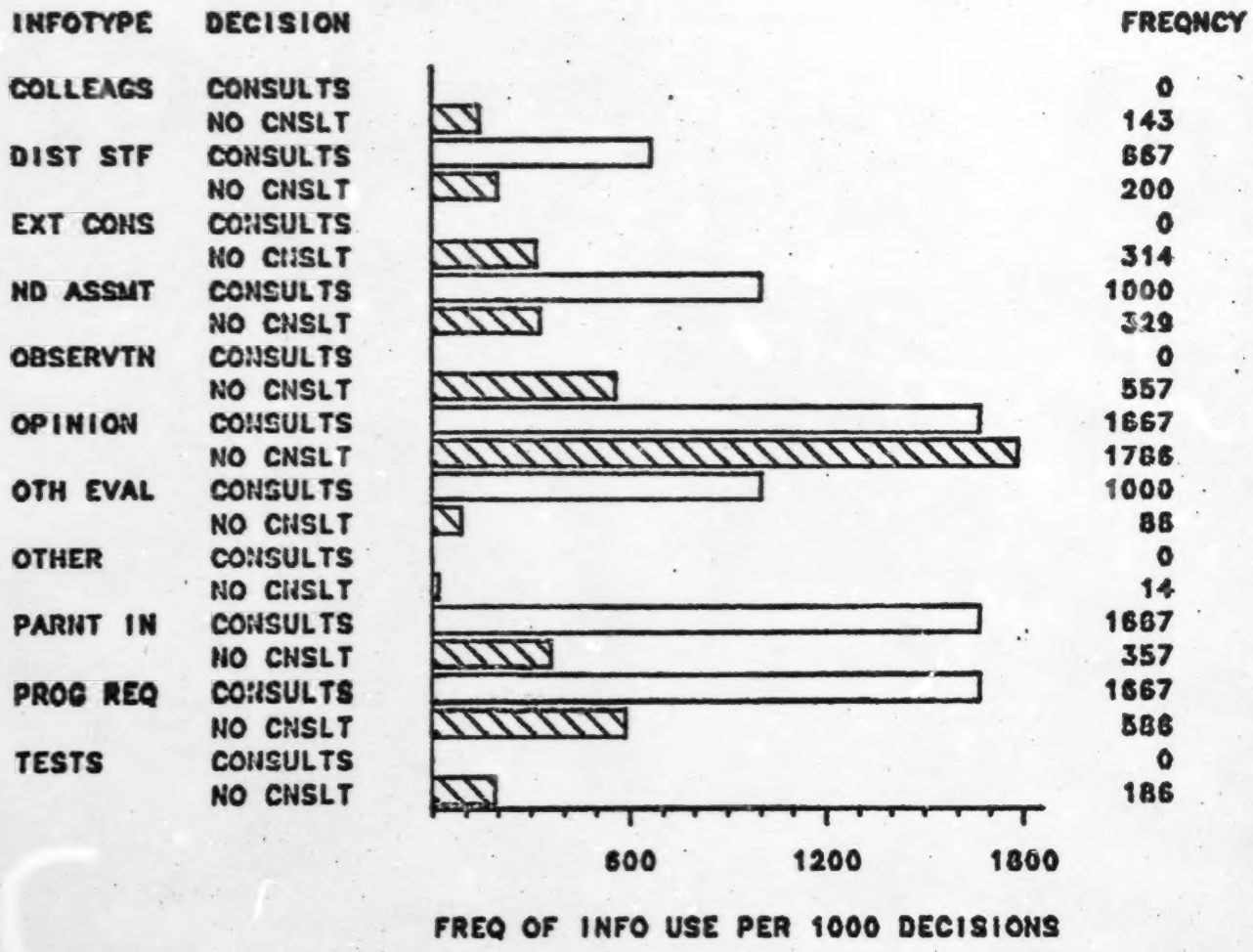

Figure F-7: Information Use, CONSULTS(N=3) vs. NO CNSLT(N=70)